



THE UNIVERSITY *of* EDINBURGH

## Edinburgh Research Explorer

# The Effects of Demography and Linkage on the Estimation of Selection and Mutation Parameters

**Citation for published version:**

Zeng, K & Charlesworth, B 2010, 'The Effects of Demography and Linkage on the Estimation of Selection and Mutation Parameters', *Genetics*, vol. 186, no. 4, pp. 1411-U537.  
<https://doi.org/10.1534/genetics.110.122150>

**Digital Object Identifier (DOI):**

[10.1534/genetics.110.122150](https://doi.org/10.1534/genetics.110.122150)

**Link:**

[Link to publication record in Edinburgh Research Explorer](#)

**Document Version:**

Publisher's PDF, also known as Version of record

**Published In:**

Genetics

**Publisher Rights Statement:**

Freely available via Pub Med.

**General rights**

Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

**Take down policy**

The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact [openaccess@ed.ac.uk](mailto:openaccess@ed.ac.uk) providing details, and we will remove access to the work immediately and investigate your claim.



# The Effects of Demography and Linkage on the Estimation of Selection and Mutation Parameters

Kai Zeng<sup>1</sup> and Brian Charlesworth

*Institute of Evolutionary Biology, School of Biological Sciences, University of Edinburgh, Edinburgh EH9 3JT, United Kingdom*

Manuscript received August 13, 2010

Accepted for publication September 27, 2010

## ABSTRACT

We explore the effects of demography and linkage on a maximum-likelihood (ML) method for estimating selection and mutation parameters in a reversible mutation model. This method assumes free recombination between sites and a randomly mating population of constant size and uses information from both polymorphic and monomorphic sites in the sample. Two likelihood-ratio test statistics were constructed under this ML framework:  $LRT_\gamma$  for detecting selection and  $LRT_\kappa$  for detecting mutational bias. By carrying out extensive simulations, we obtain the following results. When mutations are neutral and population size is constant,  $LRT_\gamma$  and  $LRT_\kappa$  follow a chi-square distribution with 1 d.f. regardless of the level of linkage, as long as the mutation rate is not very high. In addition,  $LRT_\gamma$  and  $LRT_\kappa$  are relatively insensitive to demographic effects and selection at linked sites. We find that the ML estimators of the selection and mutation parameters are usually approximately unbiased and that  $LRT_\kappa$  usually has good power to detect mutational bias. Finally, with a recombination rate that is typical for *Drosophila*,  $LRT_\gamma$  has good power to detect weak selection acting on synonymous sites. These results suggest that the method should be useful under many different circumstances.

THE strength of natural selection, the rate of mutation, and the intensity of genetic drift are key players in evolution, determining patterns of diversity within species and patterns of divergence between species (KIMURA 1983; GILLESPIE 1991). Quantifying the relative contributions of these forces to DNA and protein sequences is a central topic in modern evolutionary research (LI 1997; CHARLESWORTH and CHARLESWORTH 2010). At the population genetics level, they are often studied using either the infinite sites model (KIMURA 1971) or the reversible mutation model (WRIGHT 1931, 1949). The standard version of both models allows only two variants at each nucleotide site in the genome. The infinite sites model assumes that mutation is unidirectional, from the ancestral state at a fixed site to the derived states, with no mutational events allowed at segregating sites; the reversible mutation model allows bidirectional mutation between the two alleles, which can occur even at segregating sites. The latter model is arguably more realistic. Nonetheless, these two models are intrinsically related to each other, with the infinite sites model approximating the reversible mutation model when the product of the effective population size and the mutation rate is suf-

ficiently small (KONDRASHOV 1995; DESAI and PLOTKIN 2008; see also [supporting information, File S1](#)).

The infinite sites assumption has formed the basis of many methods for estimating mutational and/or selection parameters from polymorphism data collected within a species (SAWYER and HARTL 1992; AKASHI and SCHAEFFER 1997; BUSTAMANTE *et al.* 2002; SAWYER *et al.* 2003, 2007; EYRE-WALKER *et al.* 2006; KEIGHTLEY and EYRE-WALKER 2007; BOYKO *et al.* 2008). We refer to this class of methods as Poisson random field (PRF) methods (SAWYER and HARTL 1992). In most but not all applications, they require the polarization of ancestral and derived states at each polymorphic site using an outgroup species. Because modeling linkage between sites is theoretically and computationally extremely challenging (GRIFFITHS and MARJORAM 1996; NEUHAUSER and KRONE 1997; STUMPF and McVEAN 2003), linkage equilibrium between segregating sites is usually assumed. Simulations have shown that, when this assumption is violated, the PRF methods suffer from high false positive rates with respect to detecting selection (see Table 2 in BUSTAMANTE *et al.* 2001). To solve this problem, a method that requires estimation of recombination rates from the data and determination of significance using computer simulations has been developed (ZHU and BUSTAMANTE 2005).

The reversible mutation (RM) model has typically been used to study weak selection acting on synonymous codons (*i.e.*, codon usage bias), usually assuming that the scaled mutation rate is so low that the infinite sites

Supporting information is available online at <http://www.genetics.org/cgi/content/full/genetics.110.122150/DC1>.

<sup>1</sup>Corresponding author: Institute of Evolutionary Biology, School of Biological Sciences, University of Edinburgh, Ashworth Laboratories, King's Bldgs., W. Mains Rd., Edinburgh EH9 3JT, United Kingdom.  
E-mail: kai.zeng.cn@gmail.com

assumption is approximately valid (LI 1987; BULMER 1991; McVEAN and CHARLESWORTH 1999). But this class of methods does not necessarily require polarization of mutations. Instead, it considers several predefined allelic classes (typically two; see ZENG 2010 for a multi-allele version of the model), such as preferred and unpreferred codons at synonymous sites (LI 1987; BULMER 1991; McVEAN and CHARLESWORTH 1999) or AT *vs.* GC base pairs in noncoding regions (GALTIER *et al.* 2006; HADDRILL and CHARLESWORTH 2008; ZENG and CHARLESWORTH 2010). Under this framework, several maximum-likelihood (ML) inference methods have recently been proposed (MASIDE *et al.* 2004; CUTTER and CHARLESWORTH 2006; GALTIER *et al.* 2006; ZENG and CHARLESWORTH 2009; ZENG 2010); we refer to these methods as the RM methods. These methods also make the assumption of no linkage disequilibrium.

Given that the PRF methods are sensitive to the violation of the free-recombination assumption and that they are closely related to the RM methods, it is natural to ask whether the RM methods also tend to give false positive results in the presence of linkage. The answer to this question has important implications for the validity of the inferences that selection is acting on synonymous sites, as found in studies of *Caenorhabditis* (CUTTER and CHARLESWORTH 2006; CUTTER 2008) and *Drosophila* (MASIDE *et al.* 2004; BARTOLOMÉ *et al.* 2005; COMERON and GUTHRIE 2005; ZENG and CHARLESWORTH 2009, 2010), and that GC base pairs are favored over AT base pairs by selection or biased gene conversion in non-coding sequences (GALTIER *et al.* 2006; HADDRILL and CHARLESWORTH 2008; ZENG and CHARLESWORTH 2010).

In addition to linkage, another issue that needs consideration is demography. In their simplest form, both the PRF methods and the RM methods assume that the population is randomly mating and that its size is constant over time. These assumptions are often unrealistic. It is known that population structure or changes of population size can cause false inferences of selection from PRF methods, even if evolution is strictly neutral (ZHU and BUSTAMANTE 2005); new methods have been constructed to address this problem (WILLIAMSON *et al.* 2005; KEIGHTLEY and EYRE-WALKER 2007; BOYKO *et al.* 2008; GUTENKUNST *et al.* 2009). However, it is unclear to what extent the RM methods produce false inferences of selection when the underlying demographic history of the species is ignored, especially as RM methods that incorporate population size changes have only recently been developed (ZENG and CHARLESWORTH 2009, 2010).

Finally, and more specifically relevant to the study of codon usage bias, it is important to know whether reliable inferences of selection can be obtained when applying the RM methods to synonymous sites that are embedded in a background of nonsynonymous sites that are potentially under stronger selection. It is known

that the interaction between linked selected sites, known as Hill–Robertson interference (HRI) (HILL and ROBERTSON 1966; FELSENSTEIN 1974; McVEAN and CHARLESWORTH 2000; COMERON and KREITMAN 2002; COMERON *et al.* 2008; KAISER and CHARLESWORTH 2009; SEGER *et al.* 2010), can reduce the effectiveness of selection and distort allele frequency spectra at linked sites. We thus need to know whether HRI among sites subject to reversible mutation and selection can cause the RM methods to produce false positive results.

In this study, we focus on the method of ZENG and CHARLESWORTH (2009), a version of the RM method that is flexible enough to include the effects of demographic changes. To address the questions raised above, we generated random samples from simulated populations under various models and analyzed these samples using this method. When specific underlying assumptions are violated, we ask (1) whether selection and mutation parameters can be accurately estimated and (2) whether the tests for detecting selection and mutational bias tend to be conservative or give false positive results.

## METHODS

### The method of ZENG and CHARLESWORTH (2009):

This method is based on the RM model (WRIGHT 1949; LI 1987; BULMER 1991; McVEAN and CHARLESWORTH 1999). In its simplest version, it assumes a randomly mating Wright–Fisher population of  $N$  diploid individuals (see Table 1 for notation). Two alleles,  $A_0$  and  $A_1$ , are possible at an autosomal nucleotide site. The mutation rate from  $A_0$  to  $A_1$  is  $\kappa u$ , and that in the reverse direction is  $u$ ; mutation is said to be biased when  $\kappa \neq 1$ . The fitnesses of the three genotypes,  $A_0A_0$ ,  $A_0A_1$ , and  $A_1A_1$ , are 1,  $1 - s$ , and  $1 - 2s$  (genic selection). ZENG and CHARLESWORTH (2009) showed that the equilibrium distribution of the frequency of  $A_1$  in the population can be obtained by numerically solving a linear system jointly determined by  $N$ ,  $u$ ,  $\kappa$ , and  $s$ .

For a sample of  $n$  chromosomes, each composed of  $L$  unlinked sites and taken randomly from the population, we can count the number of sites where  $A_1$  is represented  $i$  times, denoted as  $d_i$  ( $0 \leq i \leq n$ ). We often call  $\mathbf{D} = (d_0, d_1, \dots, d_n)$  the (sample) allele frequency spectrum ( $L = \sum_{i=0}^n d_i$ ). Standard diffusion theory suggests that, under weak evolutionary forces, sampling properties can be characterized by two compound parameters:  $\theta = 4Nu$  and  $\gamma = 4Ns$  (EWENS 2004, Chap. 5). Thus, under the free recombination assumption, we can write down the likelihood function of the data,  $L(\theta, \kappa, \gamma \mid \mathbf{D})$  (ZENG and CHARLESWORTH 2009). Maximum-likelihood estimates of the parameters, denoted as  $\hat{\theta}$ ,  $\hat{\kappa}$ , and  $\hat{\gamma}$ , can then be obtained numerically using the simplex algorithm (PRESS *et al.* 1992).

**TABLE 1**  
**Summary of notation**

Notation	Description
$N$	No. of breeding diploid individuals
$N_e$	Effective population size
$L$	Length (in base pairs) of a simulated region
$u$	Per-site mutation rate from $A_1$ to $A_0$
$\kappa$	Mutational bias: the mutation rate from $A_0$ to $A_1$ is $\kappa u$
$s$ ( $s_i$ )	Selection coefficient (at the $i$ th site)
$c$	Rate of crossing over per base pair
$g$	Rate of initiation of gene conversion events per base pair
$T$	Length of a gene conversion tract
$\theta$	$4Nu$
$\rho$	$4Nc$
$\phi$	$4Ng$
$\gamma$	$4Ns$
$\hat{\theta}, \hat{\kappa}, \hat{\gamma}$	Maximum-likelihood estimates (MLEs) of $\theta, \kappa, \gamma$

To detect selection, ZENG and CHARLESWORTH (2009) proposed the use of the following likelihood-ratio test statistic:

$$\text{LRT}_{\gamma} = -2 \log \frac{L(\theta, \kappa, \gamma = 0 | \mathbf{D})}{L(\theta, \kappa, \gamma | \mathbf{D})}. \quad (1a)$$

For detecting mutational bias, the following statistic was proposed:

$$\text{LRT}_{\kappa} = -2 \log \frac{L(\theta, \kappa = 1, \gamma | \mathbf{D})}{L(\theta, \kappa, \gamma | \mathbf{D})}. \quad (1b)$$

Under the free recombination assumption, the distributions of these two statistics should converge to a chi-square distribution with 1 d.f., denoted as  $\chi^2_1$ .

**Proper statistical tests:** A statistical test is referred to as a proper test if the observed rejection rate is close to the nominal significance level, which is 5% throughout this study. Statistically, we can use the Kolmogorov–Smirnov (KS) test (as implemented in the R software package) to determine whether the distribution of the  $P$ -values obtained by applying a test of interest to the simulated data (see below) is uniform on  $[0, 1]$ , which should be the case if the test is proper.

**Tajima's  $D$ :** The skewness of an observed frequency spectrum was measured by Tajima's  $D$  (TAJIMA 1989b). To see whether the observed Tajima's  $D$  values are close to the theoretical minimum, we obtained values of the relative Tajima's  $D$  (SCHAEFFER 2002) whose minimum value is  $-1$ .

To determine the level of significance for the  $D$  value calculated from a sample with  $S$  segregating sites, we first used the coalescent process to generate 15,000 random samples, each with  $S$  segregating sites, assuming con-

stant population size and complete linkage (HUDSON 1990); a  $D$  value was obtained for each of these samples. The observed  $D$  value was regarded as significant if it was smaller (or larger) than the 5th (or 95th) percentile of the simulated distribution (WALL and HUDSON 2001).

**Forward simulation algorithms:** We used a forward simulation algorithm to generate random samples. We modeled a haploid Wright–Fisher population (EWENS 2004, p. 20), *e.g.*, with 1000 chromosomes, corresponding to a diploid population size of  $N = 500$ . A small population size was used due to computational constraints. Nonetheless, the fact that sample properties are determined by the compound parameters such as  $\theta$  and  $\gamma$  under weak evolutionary forces (EWENS 2004, Chap. 5) provides the theoretical justification for “scaling down” the population size, provided that the  $\theta$ - and  $\gamma$ -values for the true population size are preserved. This rescaling method has been widely used in population genetics and has been shown to be highly effective (MCVEAN and CHARLESWORTH 2000; TACHIDA 2000; COMERON and KREITMAN 2002; KAISER and CHARLESWORTH 2009). Haploids were used to avoid the complications of extreme associative overdominance that can arise with strong selection in a small population (CHARLESWORTH and CHARLESWORTH 1997; PALSSON 2002).

Each chromosome had  $L$  sites, each with two possible variants,  $A_0$  and  $A_1$ . Let  $L_0$  and  $L_1$  be the numbers of  $A_0$  and  $A_1$  sites in a given chromosome ( $L_0 + L_1 = L$ ). It was assumed that in this chromosome, in a given generation, a Poisson number of  $A_0$  sites with mean  $\kappa u L_0$  mutated to  $A_1$  and a Poisson number of  $A_1$  sites with mean  $u L_1$  mutated to  $A_0$ . In models with nonneutral evolution, fitnesses were recalculated after adding new mutations to all the chromosomes, assuming multiplicative fitness over selected sites,

$$w = \prod_{\{\text{all selected sites}\}} w_i, \quad (2)$$

where  $w_i = 1 - s_i$  if  $A_1$  is present at the  $i$ th selected site and  $w_i = 1$  if  $A_0$  is present at this site.

In the absence of recombination, the next generation was formed by sampling with replacement from the current generation, with the chance of sampling a particular individual being proportional to its fitness. When recombination was present, to generate a new individual, two parental chromosomes were first chosen, then recombinants were constructed, and finally one of the two recombinants was randomly chosen to be retained in the new generation.

Recombination was assumed to be caused by either crossing over or gene conversion, or both. For each pair of parental chromosomes, the numbers of crossing over and gene conversion events were drawn from Poisson distributions with mean values of  $cL$  and  $gL$ , respectively.



These events were randomly placed onto the chromosome. For a gene conversion event,  $T$  sites downstream of the initiation point were converted, where  $T$  was a geometric random variable with mean  $E(T)$ . For purposes of comparison with real populations, we used values of  $cL$  and  $gL$  such that their products with  $N$  are comparable with the corresponding products with estimates of effective population size  $N_e$  for *Drosophila* populations, as described by KAISER and CHARLESWORTH (2009).

**Neutral demographic models:** Complete linkage between sites was assumed in all neutral demographic models. To examine the effects of changes of population size, we assumed that a population of  $N_1$  diploids was initially at statistical equilibrium. At time zero, the population size changed instantly to  $N_2$  and stayed constant afterward. The time since the change in population size is measured by  $\tau = t/(2N_2)$ , where  $t$  is the number of generations since the change. Similarly, a bottleneck model (THORNTON and ANDOLFATTO 2006) was also examined.

To model population structure, we used Wright's symmetric island model with  $d$  demes (reviewed in CHARLESWORTH and CHARLESWORTH 2010, pp. 317–320). Each deme had  $2N = 500$  chromosomes. Let  $m$  be the probability that an individual is a migrant. Then the equilibrium value of the widely used measure of genetic isolation between demes,  $F_{ST}$ , is given by

$$F_{ST} = \frac{1}{1 + 4Nm(d/(d-1))^2}. \quad (3)$$

Different values of  $m$  and  $d$  were used; all haplotypes were sampled from one deme (Table 5), which is likely to maximize the distortion of variant frequency spectra caused by population structure (DE and DURRETT 2007). The effects of other sampling schemes were examined in Table S3.

**Models of codon structure and selection at nonsynonymous sites:** To model codon structure, we assumed that, along the whole length of each chromosome, a pair of selected nonsynonymous sites was followed by a synonymous site (KAISER and CHARLESWORTH 2009). At the beginning of each replicate simulation, the selection intensity for each nonsynonymous site (*i.e.*,  $\gamma_i$ ) was sampled from a log-normal distribution with shape and location parameters of 3.022 and 0.0368, respectively, and was assumed to be constant throughout this replicate. The shape and location parameters correspond to the exponentials of the standard deviation and mean of  $\ln(s)$  (LOEWE and CHARLESWORTH 2006). With  $N = 500$ , this distribution gives a harmonic mean selection coefficient,  $s_h$ , such that  $Ns_h = 10$ . This corresponds approximately to the mean selection coefficient for nonsynonymous mutations that are segregating in a typical *Drosophila* population (LOEWE and CHARLESWORTH 2006; LOEWE *et al.* 2006). The log-normal distribution was used because it has

been shown to provide good fit to the data obtained in population genetic surveys (LOEWE and CHARLESWORTH 2006; LOEWE *et al.* 2006; BOYKO *et al.* 2008). Furthermore, a recent simulation study showed that the above model can accurately predict the level of synonymous diversities observed for *Drosophila* dot and neo-Y chromosomes (KAISER and CHARLESWORTH 2009). Synonymous sites may be either neutrally evolving or subject to weak selection. In the latter case, it was assumed that all synonymous sites were subject to the same selection intensity.

**Burn-in time:** Previous simulation studies showed that the time required to reach statistical equilibrium is of the order of  $1/u$  generations under the reversible mutation model (MCVEAN and CHARLESWORTH 2000; TACHIDA 2000; COMERON and KREITMAN 2002). In this study, for each parameter combination, we inspected the amount of time needed to reach equilibrium. A burn-in period of at least  $5/u$  generations was implemented before samples were taken.

**Some implementation details:** The simulation algorithms were written in the Java programming language. A chromosome was represented by an array of 64-bit-long integers (computer words), with the state of a given bit representing the state of a nucleotide site. To improve efficiency, the object-oriented features of the Java language were exploited to reuse the integer arrays as much as possible to avoid copying.

**Predicting the effects of background selection:** In the model with codon structure and selection on nonsynonymous sites, the effective population size,  $N_e$ , as estimated from the equilibrium diversity at linked, neutrally evolving sites, is reduced by selection at linked sites (KAISER and CHARLESWORTH 2009). Under the assumption that selected sites are close to deterministic mutation–selection balance (background selection; CHARLESWORTH *et al.* 1993), the expected reduction in neutral diversity can be calculated as follows. We define  $B$  as  $N_e/N$ , where  $N$  is the effective population size in the absence of selection at linked sites. According to NORDBERG *et al.* (1996), we have

$$B \approx \exp \left\{ - \sum_i \frac{u}{s_i [1 + (1 - s_i)r_i/s_i]^2} \right\}, \quad (4)$$

where  $r_i$  is the recombination rate between the focal neutral site and the  $i$ th selected site, as given by Equation 3 of LOEWE and CHARLESWORTH (2007), and  $s_i$  is the selection coefficient against carriers of a deleterious nonsynonymous mutation at this site. For a region of length  $L$ , we calculated  $B$  only for the synonymous site in the center and ignored the spatial pattern for  $B$  reported previously (LOEWE and CHARLESWORTH 2007). To take into account the log-normal distribution of  $s_i$ , we sampled an  $s_i$  value for each nonsynonymous site and calculated  $B$ . Then we

TABLE 2

The effects of linkage on the method of ZENG and CHARLESWORTH (2009) with neutrality at all sites

Input parameters				P-value of KS test (power)		Mean MLE		
$\theta$	$\kappa$	$\rho$	$\phi$	LRT $_{\gamma}$	LRT $_{\kappa}$	$\hat{\theta}$	$\hat{\kappa}$	$\hat{\gamma}$
0.005	1	0	0	0.165	0.149	0.0051	1.01	-0.03
0.02	1	0	0	0.512	0.688	0.0200	1.01	0.00
0.1	1	0	0	0.309	0.170	0.1047	1.01	0.01
0.01	2	0	0	0.692	0.000 (97.8%)	0.0100	2.01	-0.01
0.005	4	0	0	0.533	0.000 (99.8%)	0.0050	4.08	-0.01
0.02	1	0.001	0	0.542	0.550	0.0204	1.00	-0.01
0.02	1	0.02	0	0.460	0.596	0.0201	1.01	0.00
0.02	1	0	0.057	0.138	0.327	0.0201	1.01	0.00
0.02	1	0.02	0.028	0.825	0.889	0.0200	1.01	0.01

The data were generated using a neutral model with  $L = 10$  kb and  $2N = 1000$ . When gene conversion was included, we assumed  $E(T) = 352$  bp. Five hundred random samples, each composed of 15 chromosomes, were generated for each parameter combination. The samples were analyzed by the method of ZENG and CHARLESWORTH (2009). The  $P$ -values of the two likelihood-ratio tests, LRT $_{\gamma}$  and LRT $_{\kappa}$ , were examined using the Kolmogorov–Smirnov (KS) test. When the KS test detected a significant deviation, the proportion of samples where the null hypothesis for the likelihood-ratio test was rejected (*i.e.*, the power) is given in parentheses.

repeated the sampling and calculation 1000 times and used the mean value of  $B$  as the predicted reduction in  $N_e$ .

## RESULTS

**The effects of linkage on the method of Zeng and Charlesworth under neutrality in a constant-size population:** It is known that, under the infinite sites model, linkage between sites does not change the shape of the site frequency spectrum, but merely increases the variance (HUDSON 1983; BUSTAMANTE *et al.* 2001). In Figure S1, we show that the same conclusion holds under the reversible mutation model considered here.

The maximum-likelihood inference method of ZENG and CHARLESWORTH (2009) assumes free recombination between sites and uses both polymorphic and monomorphic sites to infer the parameters of the model:  $\theta$ ,  $\kappa$ , and  $\gamma$ . The extra variance induced by linkage raises concerns about the reliability of this method. To address this issue, we used our inference method to analyze random samples generated by the forward simulation algorithm (see METHODS), assuming neutrality and a constant population size, but with various levels of linkage between sites. Reassuringly, the results in Table 2 (see also Table S1) suggest that the inference method is approximately unbiased, regardless of the level of linkage. This is consistent with a previous theoretical study (WIUF 2006), which showed that likelihood methods neglecting linkage provide consistent estimators under a wide range of neutral population genetic scenarios.

ZENG and CHARLESWORTH (2009) constructed two likelihood-ratio tests, one to detect the signal of selection ( $\gamma = 0$  *vs.*  $\gamma \neq 0$ ), denoted by LRT $_{\gamma}$ , and the other to detect the presence of mutational bias ( $\kappa = 1$  *vs.*  $\kappa \neq 1$ ),

denoted by LRT $_{\kappa}$ . Standard statistical theory suggests that, if data are generated under their respective null models, the  $P$ -values of these two tests should follow a uniform distribution on the interval  $[0, 1]$ ; they are then regarded as proper tests. We used the KS test to determine whether LRT $_{\gamma}$  and LRT $_{\kappa}$  are proper tests in the presence of linkage between sites (see METHODS). Intriguingly, both LRT $_{\gamma}$  and LRT $_{\kappa}$  are proper tests over a wide range of parameter combinations (Table 2; see also Table S1). In fact, an analysis using the quantile–quantile plot (Figure S2) suggests that, even with complete linkage, the distribution of these two test statistics conforms to a chi-square distribution with 1 d.f. ( $\chi^2_1$ ). These results suggest that LRT $_{\gamma}$  and LRT $_{\kappa}$  are likely to be proper statistical tests even when the data do not conform to the assumption of free recombination. Furthermore, LRT $_{\kappa}$  seems to have good power to detect the presence of mutational biases (>97% in Table 2; see also Table S1).

The observation that linkage has only a limited influence on LRT $_{\gamma}$  and LRT $_{\kappa}$  is counterintuitive, especially when considering the dramatic increase in variance due to tight linkage (Figure S1). In Figure 1A, we plot summary statistics for the distributions of  $\hat{\theta}$ ,  $\hat{\kappa}$ , and  $\hat{\gamma}$  as functions of the rate of crossing over, for the case where  $\theta = 0.02$ ,  $\kappa = 1$ , and  $L = 10$  kb (see Figure S3 for cases with  $L = 1$  or 100 kb). Two features are of note. First, as the above results imply, all observed distributions of  $\hat{\kappa}$  and  $\hat{\gamma}$ , regardless of the level of linkage, are indistinguishable from those obtained under the free recombination assumption. On the other hand, the distribution of  $\hat{\theta}$  becomes more variable as the level of linkage between sites increases (*i.e.*, as  $\rho = 4Nr$  decreases). Thus, the results in Figure 1A suggest that the extra variance in the frequency spectrum induced by linkage between sites manifests itself mainly in the

increased variability of the distribution of  $\hat{\theta}$ . Additional simulations also suggest that, for the parameter values we have considered, with  $\rho \geq 0.05$ , the distribution of  $\hat{\theta}$  becomes indistinguishable from that observed under free recombination (results not shown).

For  $LRT_\gamma$ , some exceptions to the above conclusions have been found. For example, when  $\theta = 0.05$ ,  $\kappa = 2$ , and  $L = 10$  kb, the KS test suggests that the distribution of the  $P$ -values deviates from uniformity ( $P = 2.58 \times 10^{-13}$ ), and, at a significance level of 5%,  $LRT_\gamma$  rejects neutrality in 14.8% of the random samples. As above, we plot the distributions of  $\hat{\theta}$ ,  $\hat{\kappa}$ , and  $\hat{\gamma}$  as functions of the rate of crossing over (Figure 1B with  $L = 10$  kb; see also Figure S3). As in the case with a lower mutation rate in Figure 1A, the variance of  $\hat{\theta}$  increases sharply as  $\rho$  approaches zero. However, with a high mutation rate, the variances of  $\hat{\kappa}$  and  $\hat{\gamma}$  also increase as  $\rho$  decreases, but to a much lesser extent than that of  $\hat{\theta}$ . In fact, using a two-sample Kolmogorov–Smirnov test, the distributions of  $\hat{\kappa}$  and  $\hat{\gamma}$  under  $\rho = 0$  are significantly different from those observed under  $\rho = \infty$  (the  $P$ -values for the two tests are  $1.8 \times 10^{-6}$  and  $7.2 \times 10^{-6}$ , respectively). When a moderate level of crossing over is introduced, the effects of linkage on the distributions of  $\hat{\kappa}$  and  $\hat{\gamma}$  vanish. For example, with  $\rho = 0.001$ , the two-sample KS test suggests that the two distributions are indistinguishable from those observed under  $\rho = \infty$ , and  $LRT_\gamma$  rejects 5.8% of the random samples.

Additional simulations suggest that, with complete linkage, a high mutation rate, and  $\kappa \neq 1$ , the false positive rate of  $LRT_\gamma$  seems to be positively correlated with  $L$ , but crossing over tends to be effective in reducing false rejections (Table S1 and Figure S3). In general, we find that  $LRT_\gamma$  tends to be too liberal when these three conditions are met simultaneously: (1)  $\kappa\theta > 0.05$ , (2)  $\kappa > 1$ , and (3)  $\rho < 0.001$  (Table S1). However, this situation has limited relevance to most eukaryote species, whose scaled per-site mutation rate is usually of the order of a few percent (see Figure 1.10 of CHARLESWORTH and CHARLESWORTH 2010).

**Neutral models with population size changes or population structure:** Many methods for detecting selection assume that population size is constant over time. It is well known that violations of this assumption can make methods that rely on the infinite sites model become counterconservative (SIMONSEN *et al.* 1995; JENSEN *et al.* 2005; NIELSEN 2005; ZHU and BUSTAMANTE 2005; ZENG *et al.* 2006). In this section, we investigate the joint influence of linkage and demography on  $LRT_\gamma$  and  $LRT_\kappa$ . In the simulations, we assumed neutral evolution and complete linkage between sites.

First, we modeled a 10-fold population expansion (*i.e.*,  $N_2/N_1 = 10$ ; see METHODS). It is known theoretically that an abrupt increase in population size causes an excess of rare mutations in the sample and, consequently, negative Tajima's  $D$  values (TAJIMA 1989a; SLATKIN and HUDSON 1991). This effect renders many

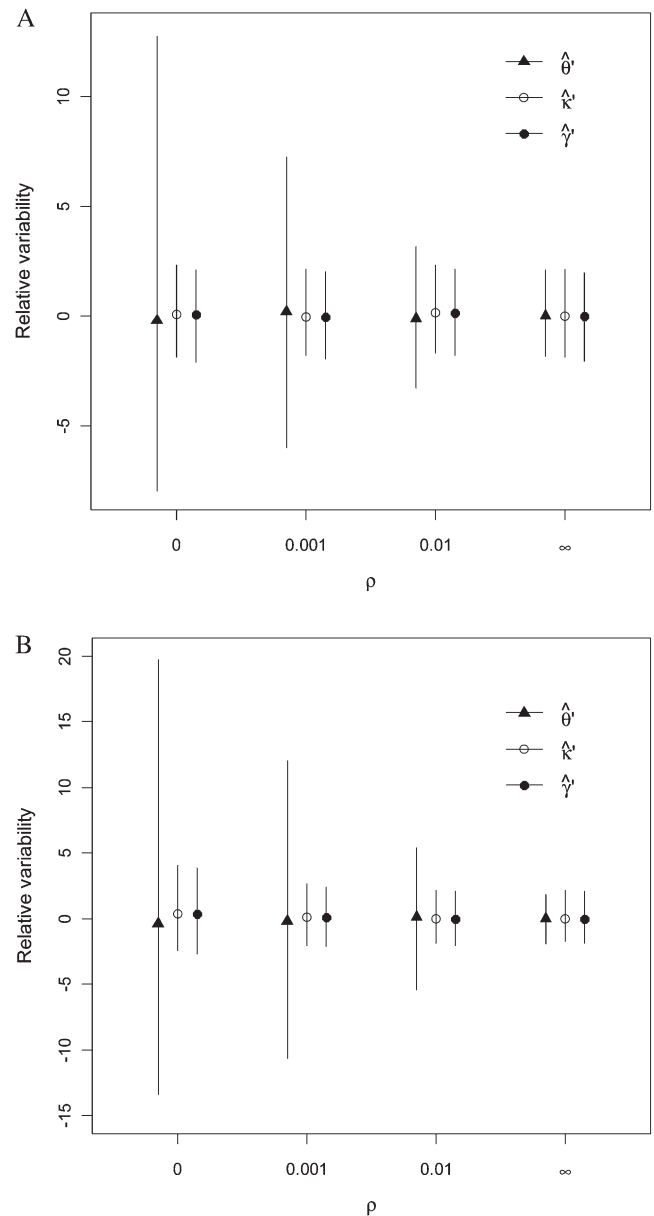


FIGURE 1.—The effects of linkage on the distributions of  $\hat{\theta}$ ,  $\hat{\kappa}$ , and  $\hat{\gamma}$ . The data were generated under the assumptions of neutrality at all sites and a constant population size. The simulated region was 10 kb in length, and the sample size was 15. The input parameters were  $\theta = 0.02$  and  $\kappa = 1$  (A) and  $\theta = 0.05$  and  $\kappa = 2$  (B). The case of  $\rho = \infty$  corresponds to the free recombination assumption. For the ease of presentation, the  $\hat{\theta}$ -values, for instance, were transformed using the formula  $\hat{\theta}' = [\hat{\theta} - E_{fr}(\hat{\theta})]/SD_{fr}(\hat{\theta})$ , where  $E_{fr}(\hat{\theta})$  and  $SD_{fr}(\hat{\theta})$  are the mean and standard deviation of  $\hat{\theta}$  obtained under free recombination, respectively; the mean of  $\hat{\theta}'$  is indicated by  $\blacktriangle$ , and the intervals between the 2.5th and 97.5th percentiles are indicated by the solid lines. Similar procedures were applied to the values of  $\hat{\kappa}$  and  $\hat{\gamma}$ .

tests (*e.g.*, Tajima's  $D$ ) counterconservative with respect to detecting purifying selection (*e.g.*, SIMONSEN *et al.* 1995; ZENG *et al.* 2006). The results in Table 3 suggest that  $LRT_\gamma$  and  $LRT_\kappa$  tend to have slightly elevated rejection rates (up to  $\sim 11\%$ ; see Table S2 for more

**TABLE 3**  
**The effects of changes of population size on  $LRT_\gamma$  and  $LRT_\kappa$**

Input: $\tau$	P-value of KS test		Observed rejection rate (%)					Mean MLE	
	$LRT_\gamma$	$LRT_\kappa$	$LRT_\gamma$	$LRT_\kappa$	$D_T$	$LRT_{\gamma 2}$	$LRT_{\kappa 2}$	$\hat{\kappa}$	$\hat{\gamma}$
$N_2/N_1 = 10, \kappa = 1$									
0.04	0.051	0.063	8.2	7.6	16.2	2.2	2.2	1.04	-0.016
0.08	$1.2 \times 10^{-5}$	$4.9 \times 10^{-5}$	11.2	11.0	28.8	4.0	4.4	1.05	-0.006
0.12	0.004	0.005	10.2	10.2	27.8	4.2	4.0	1.03	-0.005
$N_2/N_1 = 10, \kappa = 2$									
0.04	0.054	0	8.4	46.8	23.2	3.2	31.2	2.16	-0.015
0.08	$5.5 \times 10^{-4}$	0	11.2	57.4	30.8	4.4	39.8	2.15	-0.013
0.12	$6.1 \times 10^{-4}$	0	8.2	63.7	35.2	3.4	49.4	2.10	-0.009
$N_2/N_1 = 0.1, \kappa = 1$									
0.51	$2.2 \times 10^{-9}$	$18.8 \times 10^{-10}$	2.8	2.4	36.0	0.8	0.6	1.02	0.000
0.54	$6.3 \times 10^{-7}$	$3.9 \times 10^{-7}$	2.6	2.2	38.6	0.8	1.0	1.03	0.006
0.57	$8.3 \times 10^{-8}$	$2.8 \times 10^{-7}$	4.4	4.0	36.2	1.8	2.0	1.04	0.010
$N_2/N_1 = 0.1, \kappa = 2$									
0.51	0.001	0	3.8	81.2	39.4	1.6	68.4	2.09	0.008
0.54	$3.7 \times 10^{-4}$	0	4.2	80.6	36.4	1.8	67.7	2.11	0.011
0.57	0.010	0	3.8	80.2	34.8	1.0	68.2	2.13	0.017

Neutrality at all sites and complete linkage between sites were assumed. The time,  $\tau$ , since the change of population size is measured in units of  $2N_2$  generations; the values of  $\tau$  at which Tajima's  $D$  shows high rejection rates are shown. For each parameter combination, we analyzed 500 random samples of 15 chromosomes each ( $L = 10$  kb). For the expansion case ( $\kappa\theta_1 = 4N_1\kappa u = 0.002$ ), an observed value of Tajima's  $D$  ( $D_T$ ) is significant if it is smaller than the 5th percentile value generated under neutrality, using the observed number of segregating sites; for the reduction case ( $\kappa\theta_1 = 0.02$ ), a  $D_T$  is significant if it is larger than the 95th percentile.

data) under this demographic model, although these two tests seem to be more conservative than Tajima's  $D$ . In most cases, the KS test reveals that the distribution of  $P$ -values deviates significantly from the uniform distribution. Interestingly, the method seems to produce roughly unbiased estimates of  $\gamma$  and  $\kappa$ , with the  $\gamma$ -estimator behaving better than the  $\kappa$ -estimator.

Extensive empirical investigations suggest that the following statistically *ad hoc* approach is effective in solving the problem of counterconservativeness: instead of using  $\chi^2_1$ , we can perform  $LRT_\gamma$  and  $LRT_\kappa$  assuming that they follow a chi-square distribution with 2 d.f. ( $\chi^2_2$ ); we refer to these two *ad hoc* tests as  $LRT_{\gamma 2}$  and  $LRT_{\kappa 2}$ , respectively. Using data simulated under various models, we find that  $LRT_{\gamma 2}$  and  $LRT_{\kappa 2}$  are more conservative than  $LRT_\gamma$  and  $LRT_\kappa$  (see Figure S4 for more details). For example, when the data were generated under the expansion model with  $\tau = 0.04$  and  $\kappa = 1$ ,  $LRT_{\gamma 2}$  rejects 0.6, 2.2, and 5.4% of the samples at significance levels of 1, 5, and 10%, respectively;  $LRT_{\kappa 2}$  has similar properties (Table 3). When generating data using the same expansion model but with  $\kappa = 2$ , at a significance level of 5% the original  $LRT_\kappa$  rejects 46.8% of the samples, whereas the *ad hoc*  $LRT_{\kappa 2}$  rejects 31.2% (Table 3). Note that, because type I error is uncontrolled for  $LRT_\kappa$ , we cannot equate the 46.8% rejection frequency to power. On the contrary,  $LRT_{\kappa 2}$ , although conservative, has reasonable power to detect mutational bias in the

presence of recent population expansion (see Table S2 for more data).

Next, we simulated a 10-fold population size reduction (*i.e.*,  $N_2/N_1 = 0.1$ ). In contrast to the case of population expansion, a reduction in population size usually generates an excess of variants with intermediate frequencies and, consequently, positive Tajima's  $D$  values (Fu 1996). Under the assumed model, the rejection rates using Tajima's  $D$  can be as high as 40% (Table 3). For  $LRT_\gamma$  and  $LRT_\kappa$ , the KS test indicates that the distributions of these two test statistics usually do not follow a  $\chi^2_1$ -distribution. Nonetheless, all the observed rejection rates are below the 5% nominal significance level. Hence,  $LRT_\gamma$  and  $LRT_\kappa$  tend to be conservative tests following a reduction in population size, so it is unnecessary to perform the even more conservative  $LRT_{\gamma 2}$  and  $LRT_{\kappa 2}$ . When mutational bias was included,  $LRT_\kappa$  has  $\sim 80\%$  power to detect its presence, whereas  $LRT_{\kappa 2}$  has  $\sim 68\%$  power. Again, we obtain approximately unbiased estimates of  $\gamma$  and  $\kappa$ .

The third case we consider is a bottleneck model, which has been used to describe the evolution of non-African *Drosophila melanogaster* populations (THORNTON and ANDOLFATTO 2006). This model contains a brief bottleneck period, when the population size is reduced to 2.9% of the prebottleneck value, and is likely to induce false positive results with respect to detecting selection (*e.g.*, BARTON 1998). In fact, Tajima's  $D$  becomes liberal,



TABLE 4  
The effects of population bottlenecks on  $LRT_{\gamma}$  and  $LRT_{\kappa}$

Input: $\tau$	P-value of KS test		Observed rejection rate (%)					Mean MLE	
	$LRT_{\gamma}$	$LRT_{\kappa}$	$LRT_{\gamma}$	$LRT_{\kappa}$	$D_T$	$LRT_{\gamma 2}$	$LRT_{\kappa 2}$	$\hat{\kappa}$	$\hat{\gamma}$
$\kappa = 1$									
0.0084	0.144	0.205	4.8	4.6	21.4	1.6	1.8	1.21	-0.018
0.021	0.172	0.316	7.8	7.2	33.0	3.2	3.0	1.16	0.014
0.034	0.348	0.274	7.8	7.4	39.4	1.8	2.0	1.08	-0.003
0.047	0.055	0.140	7.6	7.4	41.4	3.0	3.0	1.07	-0.005
$\kappa = 2$									
0.0084	0.279	0	4.2	62.2	21.4	1.6	49.4	2.28	0.002
0.021	0.674	0	7.2	62.6	31.8	3.8	53.2	2.24	0.013
0.034	0.054	0	9.6	68.8	41.4	3.6	58.0	2.21	0.022
0.047	0.276	0	7.8	73.4	41.2	2.2	59.2	2.12	0.002

The bottleneck model of THORNTON and ANDOLFATTO (2006) for describing the evolutionary history of non-African *D. melanogaster* populations was simulated. This model assumes that the population with size  $N$  was at equilibrium until  $\tau_b = 0.044$  before the present, where time is measured in units of  $2N$  generations. At  $\tau_b$ , the population size changed instantly to  $0.029N$  and stayed constant until  $\tau_r = 0.0084$  before the present, when the population size recovered instantly to  $N$ , and stayed constant afterward.  $\tau$  is the time after the recovery of population size. The simulation assumed neutrality at all sites and complete linkage between sites. For each parameter combination, we analyzed 500 random samples of 15 chromosomes each, generated with  $4Nu = 0.015$  and  $L = 10$  kb. An observed value of Tajima's  $D$  ( $D_T$ ) is significant if it is smaller than the 5th percentile value generated under neutrality, using the observed number of segregating sites.

with up to 41.4% of the samples being rejected (Table 4). Nonetheless, the bottleneck event has only mild effects on  $LRT_{\gamma}$  and  $LRT_{\kappa}$ , and the KS test usually does not detect significant deviation from uniformity (Table 4). As under the expansion model,  $LRT_{\gamma 2}$  and  $LRT_{\kappa 2}$  tend to be very conservative. The mean value of  $\hat{\gamma}$  is close to zero, but estimates of  $\kappa$  tend to be upwardly biased, especially when the samples were taken shortly after the recovery of population size.

Finally, we consider the effects of population structure. For simplicity, we adopted a finite island model in

the simulations (see METHODS). The results are given in Table 5. First, data were generated under a two-deme model. Here, relatively low levels of geographic isolation (as measured by  $F_{ST}$ ) were assumed. None of the three tests, Tajima's  $D$ ,  $LRT_{\gamma}$ , and  $LRT_{\kappa}$  is significantly affected.  $LRT_{\kappa}$  has only moderate power ( $\sim 50\%$ ) to detect mutational bias. Next, we modeled a case with a higher level of isolation and a larger number of demes (five demes in Table 5). Under this model, Tajima's  $D$  becomes counterconservative, rejecting up to 22% of the samples. Encouragingly,  $LRT_{\gamma}$  and  $LRT_{\kappa}$  seem to be

TABLE 5  
The effects of population structure on  $LRT_{\gamma}$  and  $LRT_{\kappa}$

Input			P-value of KS test (power)		Mean MLE		
$4Nm$	$F_{ST}$	$\kappa$	$LRT_{\gamma}$	$LRT_{\kappa}$	$\hat{\kappa}$	$\hat{\gamma}$	$D_T$ (%)
Two-deme model							
1	0.2	1	0.792	0.733	1.03	-0.018	7.6
1	0.2	2	0.275	0.000 (54.0%)	2.22	0.033	6.2
0.5	0.33	1	0.067	0.007 (5.2%)	1.04	-0.009	9.6
0.5	0.33	2	0.690	0.000 (47.0%)	2.15	-0.018	9.8
Five-deme model							
1	0.39	1	0.322	0.258	1.01	-0.016	13.8
1	0.39	2	0.787	0.000 (79.8%)	2.13	0.007	14.8
0.5	0.56	1	0.342	0.296	1.03	-0.010	22.4
0.5	0.56	2	0.333	0.000 (72.6%)	2.18	0.028	18.6

Wright's island model was simulated. Neutrality at all sites and complete linkage between sites were assumed. We analyzed 500 random samples for each combination of parameter values ( $L = 10$  kb and  $\kappa\theta = 0.002$ , where  $\theta$  is defined by the number of chromosomes per deme). Each sample had 15 chromosomes, all taken from one deme. An observed value of Tajima's  $D$  is significant if it is smaller than the 5th percentile value generated under neutrality using the observed number of segregating sites; the rejection rates are given in the  $D_T$  column.

proper tests, and  $LRT_{\kappa}$  seems have reasonable power to detect mutational bias. Note that these results are based on a sampling scheme where all chromosomes are taken from one deme. Nonetheless we have obtained similar conclusions using other sampling schemes (e.g., with one chromosome taken from one deme and all the rest from another deme; Table S3).

In summary, we suggest (i) that  $LRT_{\gamma}$  and  $LRT_{\kappa}$  are relatively insensitive to the neutral demographic models we have examined, (ii) that they tend to produce approximately unbiased estimates of  $\gamma$  and  $\kappa$ , (iii) that  $LRT_{\kappa}$  seems to have reasonable power to detect mutational bias, and (iv) that the two *ad hoc* tests,  $LRT_{\gamma 2}$  and  $LRT_{\kappa 2}$ , tend to be conservative and rarely produce false positive results even when  $LRT_{\gamma}$  and  $LRT_{\kappa}$  become too liberal. The observation that the mean of  $\hat{\gamma}$  is usually close to zero, even though the frequency spectrum is distorted, is counterintuitive. In File S1, we present a numerical example based on the population expansion model, which may shed light on these seemingly paradoxical observations.

**The effects of selection at nonsynonymous sites on linked, neutrally evolving synonymous sites:** The original purpose of the method of ZENG and CHARLESWORTH (2009) was to estimate the parameters governing the evolution of synonymous sites. In this regard, the models considered in the previous sections are unrealistic, since they assume neutrality and neglect the fact that synonymous sites are linked to nonsynonymous sites in the same gene, which are under natural selection (LOEWE and CHARLESWORTH 2007). To address this issue, we conducted simulations taking into account codon structure and selection at nonsynonymous sites, but retaining the assumption of neutrality at synonymous sites (see METHODS). In particular, we assume that mutation is reversible at all sites and  $\kappa\theta = 0.02$ , which should be realistic for *D. melanogaster* (assuming  $N_e \approx 10^6$  and a mean per site mutation rate of  $5 \times 10^{-9}$ ) (KREITMAN 1983; KEIGHTLEY *et al.* 2009). We sampled the fitness effects of the selected sites from a log-normal distribution, which has been shown to provide a good fit to the data obtained from population genetic surveys (LOEWE and CHARLESWORTH 2006; BOYKO *et al.* 2008; KAISER and CHARLESWORTH 2009). The statistics obtained on synonymous sites were then analyzed to examine the performance of the method of ZENG and CHARLESWORTH (2009).

In a previous simulation study, KAISER and CHARLESWORTH (2009) concluded (i) that, with a large number of tightly linked nonsynonymous sites, the frequency spectrum at linked, neutrally evolving synonymous sites can be seriously distorted, showing extreme negative Tajima's *D* values (see also SEGER *et al.* 2010), and (ii) that gene conversion alone is ineffective in removing HRI. Here, we substantiate these conclusions by including mutational bias in the simulation and by providing distributions for the quantities of interest (Figure 2).

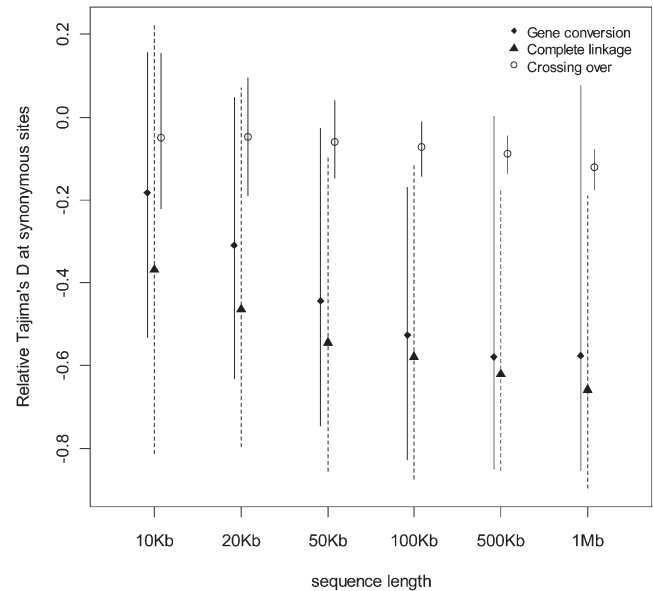


FIGURE 2.—The effects of selection at linked nonsynonymous sites on the frequency spectrum at neutral synonymous sites. Along the whole length of a chromosome, a pair of selected nonsynonymous sites is followed by a neutral synonymous site (see METHODS). We set  $\theta = 0.02$  and  $\kappa = 1$ . To see whether the observed Tajima's *D* values are close to the theoretical minimum, we present values of the relative Tajima's *D* statistic (SCHAEFFER 2002). The mean values are indicated by  $\blacktriangle$ , and the intervals between the 2.5th and 97.5th percentiles are indicated by the thin (or dashed) lines. The parameter values used in the case with gene conversion as the sole recombination mechanism were  $\phi = 0.028$  and  $E(T) = 352$  bp. For the case with crossing over alone,  $\rho = 0.04$ . For  $L < 500$  kb, the results are based on 250 random replicates; due to computational constraints, 100 and 50 replicates were performed for  $L = 500$  kb and 1 Mb, respectively.

We highlight several observations. First, incorporating mutational bias into the simulation has little impact on the frequency spectrum at synonymous sites (Figure S5). Second, the ineffectiveness of gene conversion in removing HRI is clearly demonstrated by the fact that the observed distribution of Tajima's *D* under gene conversion overlaps substantially with that observed under complete linkage (Figure 2), although gene conversion was assumed to have occurred at a rate that is realistic for *D. melanogaster* (HILLIKER *et al.* 1994; LOEWE and CHARLESWORTH 2007). In contrast, with a rate of crossing over of  $\rho = 0.04/\text{bp}$ , which should also be realistic for *D. melanogaster* [assuming  $N_e \approx 10^6$  and a mean per-site rate of crossing over of  $10^{-8}$  (HEY and KLIMAN 2002)], the distribution of Tajima's *D* is much closer to the neutral expectation of zero. As noted previously, the effects of selection at linked sites on the frequency spectrum at synonymous sites become important when there are a large number of linked selected sites. In fact, with  $>20$  kb of tightly linked sites, the 97.5th percentile of the distribution of Tajima's *D* barely overlaps with the neutral expectation of zero (Figure 2).

TABLE 6

The effects of selection at linked nonsynonymous sites on  $LRT_\gamma$  and  $LRT_\kappa$  for neutral sites

$\kappa$	Input		<i>P</i> -value of KS test		Rejection rate (%)					Mean MLE		
	<i>L</i> (kb)	$\rho$	$LRT_\gamma$	$LRT_\kappa$	$LRT_\gamma$	$LRT_\kappa$	$D_T$	$LRT_{\gamma 2}$	$LRT_{\kappa 2}$	$\hat{\theta}$	$\hat{\kappa}$	$\hat{\gamma}$
1	10	0	0.029	0.027	7.4	7.0	18.6	3.0	2.6	0.0059	1.11	-0.001
2	10	0	0.001	0	6.8	30.8	12.0	3.0	15.2	0.0029	2.34	-0.008
1	50	0	0.002	0.002	10.0	9.2	42.6	2.8	3.0	0.0023	1.05	-0.014
2	50	0	$4.0 \times 10^{-4}$	0	9.9	53.6	38.6	4.6	36.6	0.0011	2.20	0.012
1	100	0	$3.2 \times 10^{-4}$	$9.5 \times 10^{-4}$	9.4	9.2	49.6	3.0	2.6	0.0016	1.02	-0.017
2	100	0	$3.8 \times 10^{-4}$	0.000	8.0	63.2	48.8	2.6	50.4	0.0008	2.13	0.005
1	10	0.008	0.066	0.044	6.4	6.2	0.4	2.0	1.8	0.0088	1.06	-0.011
1	50	0.008	0.012	0.014	8.6	8.8	0.2	4.0	3.6	0.0048	1.02	-0.003
1	10	0.012	0.971	0.827	5.0	5.6	0.0	1.6	1.8	0.0098	1.05	-0.007
1	50	0.012	0.273	0.170	7.8	8.2	0.0	1.8	1.4	0.0060	1.02	0.002

The simulation procedures were identical to those used to obtain Figure 2. Synonymous sites were assumed to be neutral. We set  $\kappa\theta = 0.02$ . The results are based on 500 random samples of 15 chromosomes each.

To examine the details of the frequency spectra for derived alleles, we used the forward simulation algorithm to obtain the unique genealogy that relates all extant haplotypes in the absence of recombination. For each segregating site in the sample, we determined ancestral and derived alleles by comparing the extant sequences with the ancestral sequence at the most recent common ancestor for this sample. In Figure S6, the spectrum observed at synonymous sites is shown and compared to the neutral expectation. We observe an excess of low-frequency variants and a deficit of intermediate- and high-frequency ones. This is in line with the prediction under the standard background selection theory with weak selection (CHARLESWORTH *et al.* 1995; FU 1997; GORDO *et al.* 2002).

In the absence of recombination, Tajima's *D* is counterconservative when applied to synonymous sites, with observed rejection rates up to 74% (Table 6 and Table S4). For  $LRT_\gamma$  and  $LRT_\kappa$ , their *P*-values are no longer uniformly distributed (Table 6), and the observed rejection rates are slightly elevated (<14%; Table S4). As in the case of population expansion,  $LRT_{\gamma 2}$  and  $LRT_{\kappa 2}$  tend to be conservative, with all observed rejection rates <5% (Table 6, Table S4, and Figure S7). Note that, because type I error is uncontrolled, power is undefined for  $LRT_\kappa$  in the case of no recombination. On the other hand, using  $LRT_{\kappa 2}$ , we still have some power to detect mutational bias: 15.2% for  $L = 10$  kb, 36.6% for  $L = 50$  kb, and 50.4% for  $L = 100$  kb (Table 6).

Additional simulations suggest that a relatively high level of crossing over is necessary to reduce the effects of selection at linked sites on  $LRT_\gamma$  and  $LRT_\kappa$ . For example, with  $L = 10$  kb,  $\rho > 0.01$  is necessary (Table 6). Note that, since the simulation model does not include any noncoding regions, the rate of crossing over needed to alleviate the effects of selection at linked sites is likely to be overestimated. Nonetheless caution should be exercised when applying  $LRT_\gamma$  and  $LRT_\kappa$

to genomic regions where crossing over is relatively infrequent; for these regions,  $LRT_{\gamma 2}$  and  $LRT_{\kappa 2}$  are probably more reliable.

The mean ML estimates for  $\gamma$  tend to be unbiased, whereas those for  $\kappa$  seem slightly upwardly biased when recombination is infrequent (Table 6). In all parameter combinations, the mean  $\hat{\theta}$ -values are severalfold lower than the input values (Table 6), and cases with larger *L* values tend to have smaller  $\hat{\theta}$ -values. This reflects the well-known fact that selection at linked sites reduces the level of neutral diversity, reflecting a reduction in  $N_e$  (MCVEAN and CHARLESWORTH 2000; COMERON *et al.* 2008; KAISER and CHARLESWORTH 2009).

**The power to detect selection at synonymous sites in the presence of linked nonsynonymous sites:** In this section, we investigate whether  $LRT_\gamma$  has any power to detect natural selection at synonymous sites when there is strong HRI caused by linked nonsynonymous sites. To this end, we introduced weak selection on synonymous sites into the simulation model described in the previous section. Here, the selection coefficient, *s*, against a deleterious mutation at a synonymous site was assumed to be constant across the entire simulated region. The region size was set to 50 kb and the input scaled intensity of selection at synonymous sites was  $\gamma = 2$  (Table 7); very similar results were obtained when the size was 100 kb (results not shown). HRI tends to reduce  $N_e$  as measured by equilibrium diversity at neutral sites, especially when there are a large number of linked selected sites (MCVEAN and CHARLESWORTH 2000; COMERON *et al.* 2008; KAISER and CHARLESWORTH 2009). As a consequence, the weak selective differences between the two variants at synonymous sites may be invisible to natural selection, due to their very small  $N_e$ s (STEPHAN *et al.* 1999).

In the case of complete linkage, with the selection and mutation parameters for nonsynonymous sites used here, the observed diversity at neutral synonymous sites is

TABLE 7

The power to detect selection at synonymous sites in the presence of linked selected nonsynonymous sites

Input		Power (%)		Mean MLE			Prediction: <sup>a</sup>
$\rho$	$\phi$	LRT <sub><math>\gamma</math></sub>	LRT <sub><math>\gamma_2</math></sub>	$\hat{\theta}$	$\hat{\kappa}$	$\hat{\gamma}$	$\gamma$
0	0	9.0	3.2	0.0023	1.06	0.088	$3.2 \times 10^{-7}$
0	0.028	17.0	6.2	0.0027	1.08	0.170	0.009
0.004	0	16.2	6.0	0.0036	1.04	0.215	0.046
0.008	0	45.0	32.6	0.0047	1.07	0.366	0.191
0.02	0	97.4	93.6	0.0077	1.06	0.681	0.639
0.04	0	100.0	100.0	0.0107	1.08	1.044	1.070
0.04	0.028	100.0	100.0	0.0118	1.06	1.161	1.210

The simulation method was similar to that used in Table 6, with the additional feature that selection was acting on synonymous sites with intensity  $\gamma = 2$ . In all cases,  $L = 50$  kb,  $\theta = 0.02$ , and  $\kappa = 1$ . For each parameter combination, the results were based on analyzing 500 random samples of 15 chromosomes. Note that, for cases with  $\rho < 0.01$ , the power of LRT <sub>$\gamma$</sub>  should be interpreted with caution because of inflated type I error rates (Table 6).

<sup>a</sup> $\gamma$ -values predicted using the background selection formula for  $N_e$  (Equation 4).

0.0015,  $\sim 13$ -fold lower than the input value of 0.02 (Table 6). This is consistent with observed levels of diversity on the *Drosophila* dot chromosome (CHARLESWORTH *et al.* 2010). Hence, for synonymous sites under selection,  $4N_e s$  predicted from the  $N_e$  value corresponding to the neutral diversity level is equal to  $\frac{2}{13} \approx 0.15$ . The mean value of  $\hat{\gamma}$  is 0.088 (Table 7), somewhat lower than this predicted value (see DISCUSSION). Not surprisingly, with such a small  $4N_e s$ , neither LRT <sub>$\gamma$</sub>  nor LRT <sub>$\gamma_2$</sub>  has any power to detect selection (Table 7). Introducing gene conversion does not significantly increase the power of the two tests. These results imply that it will be hard to detect selection on synonymous sites in parts of the genome with extremely reduced recombination (*e.g.*, the dot chromosome in *Drosophila*), consistent with the evidence that selection on codon usage is greatly reduced in these regions (reviewed by CHARLESWORTH *et al.* 2010).

The power of LRT <sub>$\gamma$</sub>  starts to increase as the rate of crossing over becomes higher (Table 7). When  $\rho = 0.02$ , about half of the typical value of 0.04 for *D. melanogaster*, LRT <sub>$\gamma$</sub>  achieves virtually 100% power in detecting selection. Nonetheless, it should be noted that, for cases with  $\rho < 0.01$ , the power of LRT <sub>$\gamma$</sub>  should be interpreted with caution because of the problem of inflated type I error rates reported in the previous section. Encouragingly, the more conservative LRT <sub>$\gamma_2$</sub>  has only slightly reduced power compared to LRT <sub>$\gamma$</sub>  (*e.g.*, 93.6% *vs.* 97.4% for  $\rho = 0.02$ ). Further, LRT <sub>$\gamma_2$</sub>  can be applied to regions with reduced recombination, where LRT <sub>$\gamma$</sub>  is counter-conservative (*e.g.*, when  $\rho = 0.008$ ).

In all cases, the estimates of  $\kappa$  are approximately unbiased (Table 7). On the other hand, with higher levels of recombination the mean values of  $\hat{\theta}$  and  $\hat{\gamma}$

becomes closer to the input values of 0.02 and 2, although even under the highest level of recombination ( $\rho = 0.04$  and  $\phi = 0.028$ ) the mean estimates of these two parameters are still lower than the input values, suggesting that the effects of selection at linked sites have not been completely removed. This reduction is due to the action of background selection (CHARLESWORTH *et al.* 1993; HUDSON and KAPLAN 1994). In fact, using the theory developed by NORDBORG *et al.* (1996) and LOEWE and CHARLESWORTH (2007), for  $\rho = 0.02$  and 0.04, the predicted effect on  $N_e$  at synonymous sites due to selection at linked nonsynonymous sites is  $N_e/N = 0.32$  and 0.54, respectively (see Equation 4); the expected  $4N_e s$  values for linked weakly selected sites are therefore 0.64 and 1.07, which are very close to the mean values of  $\hat{\gamma}$  given in Table 7, in agreement with the results of STEPHAN *et al.* (1999) for one strongly selected and one weakly selected site. However, when recombination rates are very low, the background selection formula tends to overestimate the reduction in  $N_e$  and  $\gamma$  (Table 7) for reasons given in the DISCUSSION.

If the above conclusions are correct, then the previous estimate of  $\gamma = 1.03$  (ZENG and CHARLESWORTH 2009), obtained by analyzing twofold degenerate synonymous codons located in normally recombining regions in the *D. melanogaster* genome ( $c > 1.15$  cM/Mb; *i.e.*,  $\rho > 0.04$ ), is likely to be close to the actual intensity of selection acting on these sites.

## DISCUSSION

**Properties of the methods based on the reversible mutation model:** In a series of articles (CUTTER and CHARLESWORTH 2006; ZENG and CHARLESWORTH 2009, 2010; ZENG 2010), we proposed several methods for inferring the selection and mutation parameters using the reversible mutation model, which are especially suitable for detecting selection on synonymous mutations or on GC *vs.* AT variants in noncoding sequences. These methods have the following features: (1) They do not require the use of outgroup sequences to infer ancestral states and are thus free from the potential problem of false inference caused by misidentification of ancestral states (BAUDRY and DEPAULIS 2003; HERNANDEZ *et al.* 2007); (2) they allow simultaneous estimation of selection and mutational bias parameters from population genetic data; and (3) they can be extended to account for the effects of past changes of population size.

We have further shown here that the method of ZENG and CHARLESWORTH (2009), a representative of the class of methods cited above, is fairly insensitive to the violation of the free recombination assumption. Additionally, this method seems to be relatively insensitive to demographic changes, population structure, and selection at linked sites, compared to the commonly used



Tajima's  $D$  and related tests for detecting skews in the site frequency spectrum (reviewed by CHARLESWORTH and CHARLESWORTH 2010, pp. 287–291). In cases where our likelihood-ratio tests for detecting selection and mutational bias are counterconservative, we find that the *ad hoc* approach of performing the tests with a chi-square distribution with 2 d.f. can effectively solve the problem of inflated type I error rates, while retaining reasonable power to detect selection and mutational bias (Tables 3 and 6).

We have paid particular attention to the effect of selection at linked sites on the behavior of neutral or weakly selected sites. Consistent with earlier work (KAISER and CHARLESWORTH 2009; SEGER *et al.* 2010), we found that selection on nonsynonymous sites causes a large skew in the site frequency spectrum at neutral sites when the recombination rate is very low, with very negative values of Tajima's  $D$ , many of which approach the minimal possible value (Figure 2). This pattern could easily be mistaken for the effect of a selective sweep, as discussed by KAISER and CHARLESWORTH (2009).

In addition, as mentioned in the description of the results shown in Tables 6 and 7, when sites subject to weak purifying selection, such as synonymous or noncoding sites, are embedded in a large low-recombination region containing many strongly selected sites, the estimated scaled intensity of selection,  $\gamma = 4N_e s$ , for the weakly selected sites (Table 7) is reduced by considerably more than that predicted from the reduction in  $N_e$  caused by the strongly selected (nonsynonymous) sites, as estimated from diversity at neutral sites (see the last section of RESULTS). This is probably caused by the additional HRI among the weakly selected sites themselves (McVEAN and CHARLESWORTH 2000; COMERON and KREITMAN 2002; COMERON *et al.* 2008).

For recombination rates that are typical of normally recombining regions of the *Drosophila* or human genomes, in the presence of the strongly selected sites, there is still a reduction in  $N_e$  for neutral synonymous sites and in  $\gamma$  for weakly selected sites, which is well predicted by the background selection equation (Equation 4). But with low levels of recombination, this equation tends to overestimate the reduction in  $N_e$  and  $\gamma$  (Table 7). This phenomenon was noted previously (KAISER and CHARLESWORTH 2009) and reflects HRI between the nonsynonymous sites in regions with very low recombination rates, which weakens the effective strength of selection acting on them. As a result, the deterministic mutation–selection model on which standard background selection theory is based does not apply.

It should be noted that we have not attempted to construct a realistic model of genome structure for normally recombining genomic regions; the results shown in Tables 6 and 7 make no allowance for noncoding sequences either within or between genes, which are common in eukaryotes and can potentially serve as

spacers reducing HRI (COMERON and KREITMAN 2002), although there is evidence that noncoding regions are themselves under selection (*e.g.*, ANDOLFATTO 2005; HADDRILL *et al.* 2005). The effects of HRI that we have described here are probably overestimates of the true effects for genes in normally recombining regions. Further work is needed to determine the properties of more realistic models.

**Differences between the PRF methods and the RM methods:** The key difference between the PRF methods and the RM methods is that with PRF alleles are typically classified as ancestral and derived using outgroup sequences, whereas with RM alleles are assigned to several predetermined classes (typically two; see ZENG 2010 for a multiallele model) that are potentially selectively different. Examples of the latter include preferred and unpreferred synonymous codons (LI 1987; BULMER 1991; McVEAN and CHARLESWORTH 1999) or AT and GC in noncoding regions (GALTIER *et al.* 2006; HADDRILL and CHARLESWORTH 2008; ZENG and CHARLESWORTH 2010). The need to define selectively different allele classes makes it hard to apply the RM approach to such problems as the distribution of fitness effects of nonsynonymous mutations (EYRE-WALKER *et al.* 2006; KEIGHTLEY and EYRE-WALKER 2007; BOYKO *et al.* 2008). Further work is needed to determine whether the desirable properties of the RM methods can be used for this purpose.

A question that is raised by our results is the fact that close linkage does not seem to produce a high variance in the estimates of  $\gamma$  and  $\kappa$  from the RM method (Figure 1), in contrast to the well-known effect of close linkage on the variance of estimates of  $\theta$  (HUDSON 1983, 1990). The intuitive reason for this is that the variance of the estimate of  $\theta$  reflects the large stochastic variation in the size of gene genealogies; linked sites have a high covariance of tree sizes, which inflates the variance of the mean tree size across sites compared with that of a set of independent genealogies (PLUZHNIKOV and DONNELLY 1996; McVEAN 2002). The associated variability in the shapes of trees will also generate more variability in the frequencies of derived *vs.* ancestral variants for sets of closely linked sites compared with independent sites, resulting in a higher frequency of false positive results of tests for selection from PRF methods (BUSTAMANTE *et al.* 2001; ZHU and BUSTAMANTE 2005; DESAI and PLOTKIN 2008).

In contrast, the estimate of  $\gamma$  from the RM approach comes from the frequencies of favored *vs.* disfavored variants, such that a skew toward higher frequencies of favored variants indicates the action of selection or biased gene conversion. A high-frequency, putatively favorable variant could be either derived or ancestral, so that the chance of observing it is relatively independent of tree size or shape. Similarly, estimates of  $\kappa$  come mainly from the relative frequencies of sites that are fixed for the different alleles in the model (ZENG and

CHARLESWORTH 2009; ZENG 2010), which is also little affected by tree size or shape. Similar considerations may apply to the effects of population structure and population size changes. Overall, therefore, there are reasons to expect that the RM approach is more robust than the PRF method. However, it is worth emphasizing that, due to computational constraints, the results presented in this study are based on a limited number of combinations of parameter values. Therefore, they should not be overinterpreted.

We thank Andrew Rambaut for valuable discussions on various computer algorithms. We also thank Noah Rosenberg and two anonymous reviewers for helpful comments. This study made use of the resources provided by the Edinburgh Compute and Data Facility (ECDF) (<http://www.ecdf.ed.ac.uk/>). The ECDF is partially supported by the e-Science Data, Information, and Knowledge Transformation (eDIKT) initiative (<http://www.edikt.org.uk>). K.Z. acknowledges support from a Biomedical Personal Research Fellowship awarded by the Royal Society of Edinburgh and the Caledonian Research Foundation.

#### LITERATURE CITED

- AKASHI, H., and S. W. SCHAEFFER, 1997 Natural selection and the frequency distributions of "silent" DNA polymorphism in *Drosophila*. *Genetics* **146**: 295–307.
- ANDOLFATTO, P., 2005 Adaptive evolution of non-coding DNA in *Drosophila*. *Nature* **437**: 1149–1152.
- BARTOLOMÉ, C., X. MASIDE, S. YI, A. L. GRANT and B. CHARLESWORTH, 2005 Patterns of selection on synonymous and nonsynonymous variants in *Drosophila miranda*. *Genetics* **169**: 1495–1507.
- BARTON, N. H., 1998 The effect of hitch-hiking on neutral genealogies. *Genet. Res.* **72**: 123–133.
- BAUDRY, E., and F. DEPAULIS, 2003 Effect of misoriented sites on neutrality tests with outgroup. *Genetics* **165**: 1619–1622.
- BOYKO, A. R., S. H. WILLIAMSON, A. R. INDAP, J. D. DEGENHARDT, R. D. HERNANDEZ *et al.*, 2008 Assessing the evolutionary impact of amino acid mutations in the human genome. *PLoS Genet.* **4**: e1000083.
- BULMER, M., 1991 The selection-mutation-drift theory of synonymous codon usage. *Genetics* **129**: 897–907.
- BUSTAMANTE, C. D., J. WAKELEY, S. SAWYER and D. L. HARTL, 2001 Directional selection and the site-frequency spectrum. *Genetics* **159**: 1779–1788.
- BUSTAMANTE, C. D., R. NIELSEN, S. A. SAWYER, K. M. OLSEN, M. D. PURUGGANAN *et al.*, 2002 The cost of inbreeding in *Arabidopsis*. *Nature* **416**: 531–534.
- CHARLESWORTH, B., and D. CHARLESWORTH, 1997 Rapid fixation of deleterious alleles can be caused by Muller's ratchet. *Genet. Res.* **70**: 63–73.
- CHARLESWORTH, B., and D. CHARLESWORTH, 2010 *Elements of Evolutionary Genetics*. Roberts & Co., Greenwood Village, CO.
- CHARLESWORTH, B., M. T. MORGAN and D. CHARLESWORTH, 1993 The effect of deleterious mutations on neutral molecular variation. *Genetics* **134**: 1289–1303.
- CHARLESWORTH, B., A. J. BETANCOURT, V. B. KAISER and I. GORDO, 2010 Genetic recombination and molecular evolution. *Cold Spring Harbor Symp. Quant. Biol.* **74**: 177–186.
- CHARLESWORTH, D., B. CHARLESWORTH and M. T. MORGAN, 1995 The pattern of neutral molecular variation under the background selection model. *Genetics* **141**: 1619–1632.
- COMERON, J. M., and T. B. GUTHRIE, 2005 Intragenic Hill-Robertson interference influences selection intensity on synonymous mutations in *Drosophila*. *Mol. Biol. Evol.* **22**: 2519–2530.
- COMERON, J. M., and M. KREITMAN, 2002 Population, evolutionary and genomic consequences of interference selection. *Genetics* **161**: 389–410.
- COMERON, J. M., A. WILLIFORD and R. M. KLIMAN, 2008 The Hill-Robertson effect: evolutionary consequences of weak selection and linkage in finite populations. *Heredity* **100**: 19–31.
- CUTTER, A. D., 2008 Multilocus patterns of polymorphism and selection across the X chromosome of *Caenorhabditis remanei*. *Genetics* **178**: 1661–1672.
- CUTTER, A. D., and B. CHARLESWORTH, 2006 Selection intensity on preferred codons correlates with overall codon usage bias in *Caenorhabditis remanei*. *Curr. Biol.* **16**: 2053–2057.
- DE, A., and R. DURRETT, 2007 Stepping-stone spatial structure causes slow decay of linkage disequilibrium and shifts the site frequency spectrum. *Genetics* **176**: 969–981.
- DESAI, M. M., and J. B. PLOTKIN, 2008 The polymorphism frequency spectrum of finitely many sites under selection. *Genetics* **180**: 2175–2191.
- EWENS, W. J., 2004 *Mathematical Population Genetics*. Springer-Verlag, Berlin.
- EYRE-WALKER, A., M. WOOLFIT and T. PHELPS, 2006 The distribution of fitness effects of new deleterious amino acid mutations in humans. *Genetics* **173**: 891–900.
- FELSENSTEIN, J., 1974 The evolutionary advantage of recombination. *Genetics* **78**: 737–756.
- FU, Y. X., 1996 New statistical tests of neutrality for DNA samples from a population. *Genetics* **143**: 557–570.
- FU, Y. X., 1997 Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. *Genetics* **147**: 915–925.
- GALTIER, N., E. BAZIN and N. BIERNE, 2006 GC-biased segregation of noncoding polymorphisms in *Drosophila*. *Genetics* **172**: 221–228.
- GILLESPIE, J. H., 1991 *The Cause of Molecular Evolution*. Oxford University Press, Oxford.
- GORDO, I., A. NAVARRO and B. CHARLESWORTH, 2002 Muller's ratchet and the pattern of variation at a neutral locus. *Genetics* **161**: 835–848.
- GRIFFITHS, R. C., and P. MARJORAM, 1996 Ancestral inference from samples of DNA sequences with recombination. *J. Comput. Biol.* **3**: 479–502.
- GUTENKUNST, R. N., R. D. HERNANDEZ, S. H. WILLIAMSON and C. D. BUSTAMANTE, 2009 Inferring the joint demographic history of multiple populations from multidimensional SNP frequency data. *PLoS Genet.* **5**: e1000695.
- HADRILL, P. R., and B. CHARLESWORTH, 2008 Non-neutral processes drive the nucleotide composition of non-coding sequences in *Drosophila*. *Biol. Lett.* **4**: 438–441.
- HADRILL, P. R., B. CHARLESWORTH, D. L. HALLIGAN and P. ANDOLFATTO, 2005 Patterns of intron sequence evolution in *Drosophila* are dependent upon length and GC content. *Genome Biol.* **6**: R67.
- HERNANDEZ, R. D., S. H. WILLIAMSON and C. D. BUSTAMANTE, 2007 Context dependence, ancestral misidentification, and spurious signatures of natural selection. *Mol. Biol. Evol.* **24**: 1792–1800.
- HEY, J., and R. M. KLIMAN, 2002 Interactions between natural selection, recombination and gene density in the genes of *Drosophila*. *Genetics* **160**: 595–608.
- HILL, W. G., and A. ROBERTSON, 1966 The effect of linkage on limits to artificial selection. *Genet. Res.* **8**: 269–294.
- HILLIKER, A. J., G. HARAUIZ, A. G. REAUME, M. GRAY, S. H. CLARK *et al.*, 1994 Meiotic gene conversion tract length distribution within the rosy locus of *Drosophila melanogaster*. *Genetics* **137**: 1019–1026.
- HUDSON, R. R., 1983 Properties of a neutral allele model with intragenic recombination. *Theor. Popul. Biol.* **23**: 183–201.
- HUDSON, R. R., 1990 Gene genealogies and the coalescent process, pp. 1–44 in *Oxford Surveys in Evolutionary Biology*, Vol. 7, edited by D. FUTUYMA and J. ANTONOVICS. Oxford University Press, Oxford.
- HUDSON, R. R., and N. L. KAPLAN, 1994 Gene trees with background selection, pp. 140–153 in *Non-Neutral Evolution: Theories and Molecular Data*, edited by B. GOLDING. Chapman & Hall, London.
- JENSEN, J. D., Y. KIM, V. B. DUMONT, C. F. AQUADRO and C. D. BUSTAMANTE, 2005 Distinguishing between selective sweeps and demography using DNA polymorphism data. *Genetics* **170**: 1401–1410.
- KAISER, V. B., and B. CHARLESWORTH, 2009 The effects of deleterious mutations on evolution in non-recombining genomes. *Trends Genet.* **25**: 9–12.
- KEIGHTLEY, P. D., and A. EYRE-WALKER, 2007 Joint inference of the distribution of fitness effects of deleterious mutations and pop-

- ulation demography based on nucleotide polymorphism frequencies. *Genetics* **177**: 2251–2261.
- KEIGHTLEY, P. D., U. TRIVEDI, M. THOMSON, F. OLIVER, S. KUMAR *et al.*, 2009 Analysis of the genome sequences of three *Drosophila melanogaster* spontaneous mutation accumulation lines. *Genome Res.* **19**: 1195–1201.
- KIMURA, M., 1971 Theoretical foundation of population genetics at the molecular level. *Theor. Popul. Biol.* **2**: 174–208.
- KIMURA, M., 1983 *The Neutral Theory of Molecular Evolution*. Cambridge University Press, Cambridge, UK.
- KONDRASHOV, A. S., 1995 Contamination of the genome by very slightly deleterious mutations: Why have we not died 100 times over? *J. Theor. Biol.* **175**: 583–594.
- KREITMAN, M., 1983 Nucleotide polymorphism at the alcohol dehydrogenase locus of *Drosophila melanogaster*. *Nature* **304**: 412–417.
- LI, W. H., 1987 Models of nearly neutral mutations with particular implications for nonrandom usage of synonymous codons. *J. Mol. Evol.* **24**: 337–345.
- LI, W. H., 1997 *Molecular Evolution*. Sinauer Associates, Sunderland, MA.
- LOEWE, L., and B. CHARLESWORTH, 2006 Inferring the distribution of mutational effects on fitness in *Drosophila*. *Biol. Lett.* **2**: 426–430.
- LOEWE, L., and B. CHARLESWORTH, 2007 Background selection in single genes may explain patterns of codon bias. *Genetics* **175**: 1381–1393.
- LOEWE, L., B. CHARLESWORTH, C. BARTOLOME and V. NOEL, 2006 Estimating selection on nonsynonymous mutations. *Genetics* **172**: 1079–1092.
- MASIDE, X., A. W. LEE and B. CHARLESWORTH, 2004 Selection on codon usage in *Drosophila americana*. *Curr. Biol.* **14**: 150–154.
- MCVEAN, G. A., 2002 A genealogical interpretation of linkage disequilibrium. *Genetics* **162**: 987–991.
- MCVEAN, G. A. T., and B. CHARLESWORTH, 1999 A population genetic model for the evolution of synonymous codon usage: patterns and predictions. *Genet. Res.* **74**: 145–158.
- MCVEAN, G. A., and B. CHARLESWORTH, 2000 The effects of Hill-Robertson interference between weakly selected mutations on patterns of molecular evolution and variation. *Genetics* **155**: 929–944.
- NEUHAUSER, C., and S. M. KRONE, 1997 The genealogy of samples in models with selection. *Genetics* **145**: 519–534.
- NIELSEN, R., 2005 Molecular signatures of natural selection. *Annu. Rev. Genet.* **39**: 197–218.
- NORDBORG, M., B. CHARLESWORTH and D. CHARLESWORTH, 1996 The effect of recombination on background selection. *Genet. Res.* **67**: 159–174.
- PALSSON, S., 2002 Selection on a modifier of recombination rate due to linked deleterious mutations. *J. Hered.* **93**: 22–26.
- PLUZHNIKOV, A., and P. DONNELLY, 1996 Optimal sequencing strategies for surveying molecular genetic diversity. *Genetics* **144**: 1247–1262.
- PRESS, W. H., S. A. TEUKOLSKY, W. T. VETTERLING and B. P. FLANNERY, 1992 *Numerical Recipes in C: The Art of Scientific Computing*. Cambridge University Press, Cambridge, UK.
- SAWYER, S. A., and D. L. HARTL, 1992 Population genetics of polymorphism and divergence. *Genetics* **132**: 1161–1176.
- SAWYER, S. A., R. J. KULATHINAL, C. D. BUSTAMANTE and D. L. HARTL, 2003 Bayesian analysis suggests that most amino acid replacements in *Drosophila* are driven by positive selection. *J. Mol. Evol.* **57**(Suppl. 1): S154–S164.
- SAWYER, S. A., J. PARSCH, Z. ZHANG and D. L. HARTL, 2007 Prevalence of positive selection among nearly neutral amino acid replacements in *Drosophila*. *Proc. Natl. Acad. Sci. USA* **104**: 6504–6510.
- SCHAEFFER, S. W., 2002 Molecular population genetics of sequence length diversity in the *Adh* region of *Drosophila pseudoobscura*. *Genet. Res.* **80**: 163–175.
- SEGER, J., W. A. SMITH, J. J. PERRY, J. HUNN, Z. A. KALISZEWSKA *et al.*, 2010 Gene genealogies strongly distorted by weakly interfering mutations in constant environments. *Genetics* **184**: 529–545.
- SIMONSEN, K. L., G. A. CHURCHILL and C. F. AQUADRO, 1995 Properties of statistical tests of neutrality for DNA polymorphism data. *Genetics* **141**: 413–429.
- SLATKIN, M., and R. R. HUDSON, 1991 Pairwise comparisons of mitochondrial DNA sequences in stable and exponentially growing populations. *Genetics* **129**: 555–562.
- STEPHAN, W., B. CHARLESWORTH and G. MCVEAN, 1999 The effect of background selection at a single locus on weakly selected, partially linked variants. *Genet. Res.* **73**: 133–146.
- STUMPF, M. P., and G. A. MCVEAN, 2003 Estimating recombination rates from population-genetic data. *Nat. Rev. Genet.* **4**: 959–968.
- TACHIDA, H., 2000 Molecular evolution in a multisite nearly neutral mutation model. *J. Mol. Evol.* **50**: 69–81.
- TAJIMA, F., 1989a The effect of change in population size on DNA polymorphism. *Genetics* **123**: 597–601.
- TAJIMA, F., 1989b Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics* **123**: 585–595.
- THORNTON, K., and P. ANDOLFATTO, 2006 Approximate Bayesian inference reveals evidence for a recent, severe bottleneck in a Netherlands population of *Drosophila melanogaster*. *Genetics* **172**: 1607–1619.
- WALL, J. D., and R. R. HUDSON, 2001 Coalescent simulations and statistical tests of neutrality. *Mol. Biol. Evol.* **18**: 1134–1135; author reply 1136–1138.
- WILLIAMSON, S. H., R. HERNANDEZ, A. FLEDEL-ALON, L. ZHU, R. NIELSEN *et al.*, 2005 Simultaneous inference of selection and population growth from patterns of variation in the human genome. *Proc. Natl. Acad. Sci. USA* **102**: 7882–7887.
- WU, C., 2006 Consistency of estimators of population scaled parameters using composite likelihood. *J. Math. Biol.* **53**: 821–841.
- WRIGHT, S., 1931 Evolution in Mendelian populations. *Genetics* **16**: 97–159.
- WRIGHT, S., 1949 Adaptation and selection, pp. 365–389 in *Genetics, Palaeontology, and Evolution*, edited by G. JEPSON, G. SIMPSON and E. MAYR. Princeton University Press, Princeton, NJ.
- ZENG, K., 2010 A simple multi-allele model and its application to identifying preferred/unpreferred codons using polymorphism data. *Mol. Biol. Evol.* **27**: 1327–1337.
- ZENG, K., and B. CHARLESWORTH, 2009 Estimating selection intensity on synonymous codon usage in a nonequilibrium population. *Genetics* **183**: 651–662.
- ZENG, K., and B. CHARLESWORTH, 2010 Studying patterns of recent evolution at synonymous sites and intronic sites in *Drosophila melanogaster*. *J. Mol. Evol.* **70**: 116–128.
- ZENG, K., Y. X. FU, S. SHI and C. I. WU, 2006 Statistical tests for detecting positive selection by utilizing high-frequency variants. *Genetics* **174**: 1431–1439.
- ZHU, L., and C. D. BUSTAMANTE, 2005 A composite-likelihood approach for detecting directional selection from DNA sequence data. *Genetics* **170**: 1411–1421.

Communicating editor: N. A. ROSENBERG

# GENETICS

## **Supporting Information**

<http://www.genetics.org/cgi/content/full/genetics.110.122150/DC1>

## **The Effects of Demography and Linkage on the Estimation of Selection and Mutation Parameters**

**Kai Zeng and Brian Charlesworth**

Copyright © 2010 by the Genetics Society of America  
DOI: 10.1534/genetics.110.122150



**FILE S1**

We adopt the matrix-based representation of the reversible mutation model described in ZENG and CHARLESWORTH (2009). Consider a diploid Wright-Fisher population (EWENS 2004, p. 20), whose size in generation  $t$  is  $N_t$ . At an autosomal nucleotide site, two selectively indifferent variants,  $A_0$  and  $A_1$ , can occur. Mutation is reversible: the per generation mutation rate from  $A_0$  and  $A_1$  is  $\kappa u$ , and that in the reverse direction is  $u$ . Let  $f_i(t)$  be the probability that  $A_1$  is represented  $i$  times in the population at time  $t$ , and  $p_{ij}(t)$  be the probability that the number of copies of  $A_1$  changes from  $i$  in generation  $t$  to  $j$  in generation  $t+1$  (Eq. 3 in ZENG and CHARLESWORTH 2009). Standard Markov chain theory implies the following recursion relationship (KARLIN and TAYLOR 1975)

$$f_j(t+1) = \sum_{i=0}^{2N_t} f_i(t) p_{ij}(t), \quad 0 \leq j \leq 2N_{t+1}. \quad (\text{A1})$$

ZENG and CHARLESWORTH (2009) showed that  $f_i(t)$  can be decomposed into two sub-processes

$$f_i(t) = f_{1,i}(t) + f_{0,2N_t-i}(t), \quad 1 \leq i \leq 2N_t - 1 \quad (\text{A2})$$

In (A2),  $f_{1,i}(t)$  is the probability that the mutant allele  $A_1$  is currently represented  $i$  times in the population, having originated at a site fixed for  $A_0$ . The second sub-process involves mutations originating at sites fixed for  $A_1$ , such that the mutant allele  $A_0$  is currently represented  $2N_t - i$  times in the population (i.e.,  $A_1$  is represented  $i$  times). For the first sub-process, we have the following recursion formula

$$f_{1,i}(t+1) = \sum_{k=1}^{2N_t-1} f_{1,k}(t) p_{ki}(t) + f_0(t) p_{0i}(t), \quad 1 \leq i \leq 2N_{t+1} - 1. \quad (\text{A3})$$

Under the infinite sites assumption of  $\theta_t = 4N_t u \ll 1$ , (A3) can be approximated by

$$f_{1,i}^*(t+1) = \sum_{k=1}^{2N_t-1} f_{1,k}^*(t) p_{ki}^*(t) + \frac{1}{2} f_0(t) \kappa \theta_{t+1} \delta_{i,1} \quad (\text{A4})$$

where  $\delta_{i,1}$  is the Kronecker delta, which takes the value of one if  $i = 1$  and the value of zero otherwise, and

$$p_{ij}^*(t) = \binom{2N_{t+1}}{j} \left( \frac{i}{2N_t} \right)^j \left( 1 - \frac{i}{2N_t} \right)^{2N_{t+1}-j}.$$

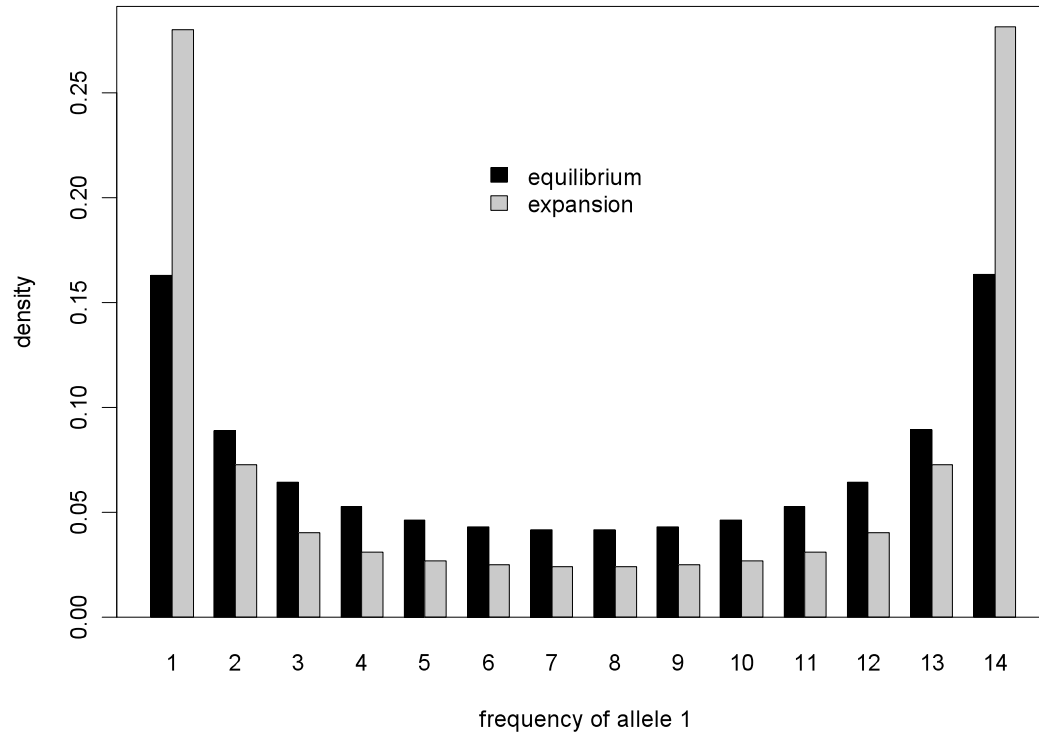
Thus, by defining

$$\begin{aligned} f_i^*(t+1) &= f_{1,i}^*(t+1) + f_{0,i}^*(t+1) \\ &= \sum_{k=1}^{2N_t-1} f_k^*(t) p_{ki}^*(t) + \frac{1}{2} \theta_{t+1} \delta_{i,1} [\kappa f_0(t) + f_{2N_t}(t)] \end{aligned} \quad (\text{A5})$$

we arrive at the recursion formula for the frequency spectrum under the infinite sites model (EVANS *et al.* 2007). Therefore, the infinite sites model can be viewed as an approximation to the reversible mutation model.

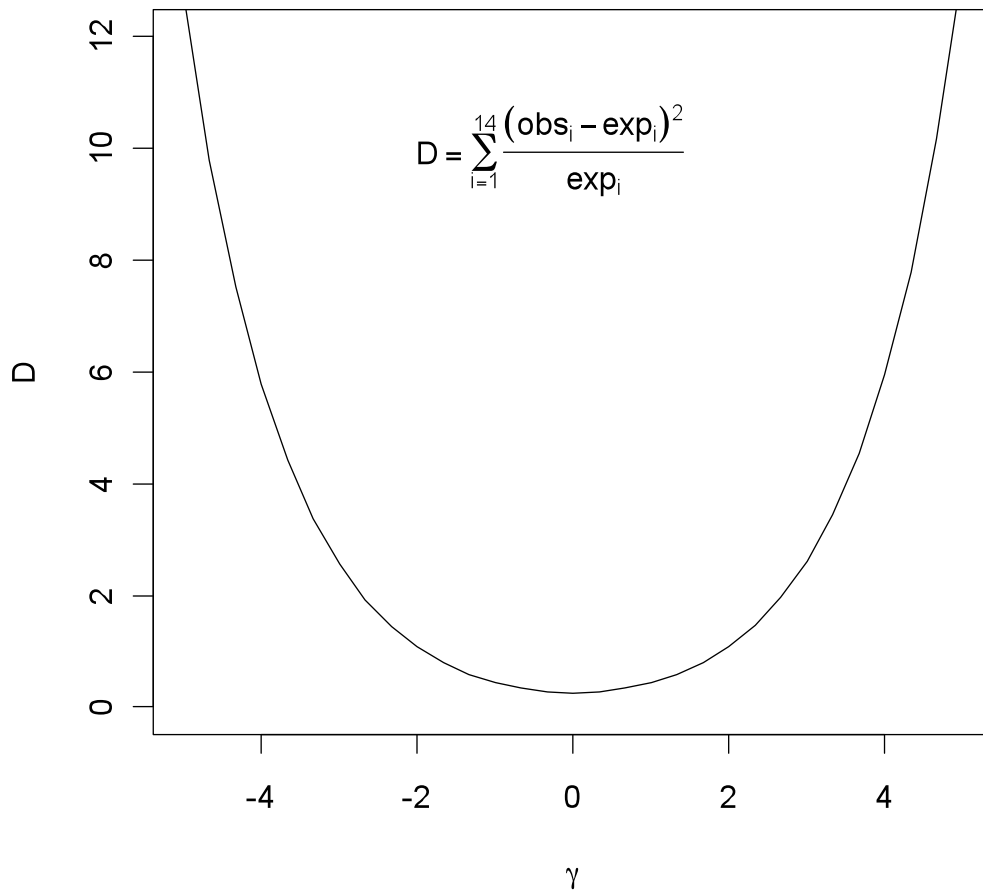
We used a population expansion model to illustrate why the method of ZENG and CHARLESWORTH (2009) tends to produce unbiased estimates of  $\gamma$ . As in Table 3, we assumed that the population was at equilibrium up to time zero, at which point its population size increased 10-fold (from  $2N_1 = 500$  to  $2N_2 = 5,000$ ) instantly and stayed constant thereafter. We used the numerical method described in PRESS *et al.* (1992) to obtain the equilibrium solutions to (A1) – (A5), in which case these formulae are independent of the time variable  $t$ . Then, these solutions were used as the initial conditions to iterate (A1) – (A5) to obtain the frequency spectra at  $\tau = t / (2N_2) = 0.04$  after the expansion event.

Using diffusion theory, McVEAN and CHARLESWORTH (1999) have shown that when the mutation rate is low (i.e.,  $\theta \ll 1$ ) the shape of the spectrum at statistical equilibrium is determined solely by  $\gamma$  (i.e., it is independent of the mutational parameters), and that the spectrum is U-shaped under neutrality (see Figure ST1 below). After the expansion, our numerical calculations suggest that that, although mutation is biased with  $\kappa = 2$ , the spectrum is still approximately U-shaped (Figure ST1). However, compared to the equilibrium case, we observe an excess of low frequency variants and a deficit of intermediate frequency variants. Because most information about selection probably comes from the segregating sites in the sample (ZENG 2010), we investigated which equilibrium spectrum (determined solely by  $\gamma$ ) is closest to the spectrum we observe after the expansion. The results are presented in Figure ST2 below. Despite the differences between the two spectra present in Figure ST1, the equilibrium spectrum obtained under  $\gamma = 0$  provides the best fit to the non-equilibrium spectrum obtained after the expansion. This may explain why the method of ZENG and CHARLESWORTH (2009) tend to give unbiased estimates of  $\gamma$ , although the assumption of statistical equilibrium has been violated by introducing changes in population size.

**Figure ST1.**

**Frequency spectra obtained under equilibrium and population expansion.** Evolution was assumed to be neutral.

The population expansion model with  $\kappa = 2$  used in Table 3 was considered. The time since the change in population size was  $\tau = 0.04$ .

**Figure ST2.**

**The distance between the non-equilibrium spectrum given in Figure ST1 and the equilibrium spectra obtained under various  $\gamma$  values.** For each value of  $\gamma$ , the equilibrium spectrum was obtained by the method of ZENG and CHARLESWORTH (2009), and its distance from the non-equilibrium spectrum obtained at  $\tau = 0.04$  after the expansion (Figure ST1) was calculated using the formula given in the figure, where  $\text{obs}_i$  is the observed frequency of sites at which allele  $A_1$  is represented  $i$  times in the sample at  $\tau = 0.04$ , and  $\text{exp}_i$  is that expected under the equilibrium model.



**REFERENCE**

- EVANS, S. N., Y. SHVETS and M. SLATKIN, 2007 Non-equilibrium theory of the allele frequency spectrum. *Theor Popul Biol* **71**: 109-119.
- EWENS, W.J., 2004 *Mathematical Population Genetics*. Springer-Verlag, Berlin.
- KARLIN, S., and H. TAYLOR, 1975 *A First Course in Stochastic Processes*. Academic Press.
- MCVEAN, G. A. T., and B. CHARLESWORTH, 1999 A population genetic model for the evolution of synonymous codon usage: patterns and predictions. *Genetical Research* **74**: 145-158.
- PRESS, W. H., S. A. TEUKOLSKY, W. T. VETTERLING and B. P. FLANNERY, 1992 *Numerical Recipes in C: The Art of Scientific Computing*. Cambridge University Press, Cambridge.
- ZENG, K., 2010 A simple multi-allele model and its application to identifying preferred/unpreferred codons using polymorphism data. *Mol Biol Evol* **27**: 1327-1337.
- ZENG, K., and B. CHARLESWORTH, 2009 Estimating selection intensity on synonymous codon usage in a nonequilibrium population. *Genetics* **183**: 651-662, 651SI-623SI.

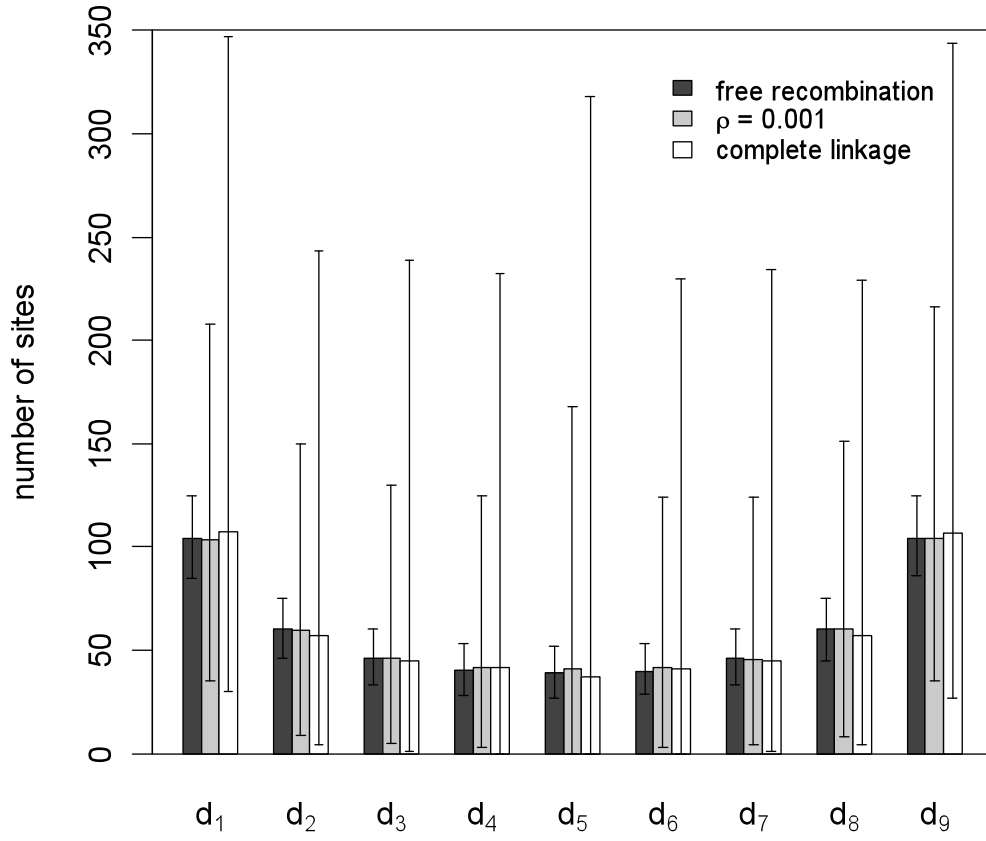


FIGURE S1.—The effects of linkage on the allele-frequency spectrum under neutrality. We simulated a region of 10Kb sites. The mutation parameter values used were  $\theta = 0.02$  and  $\kappa = 1$  (see Table 1 for definitions of the parameters used here). Three levels of linkage were considered: (i) free recombination; (ii)  $\rho = 0.001$ ; (iii) complete linkage ( $\rho = 0$ ). One thousand random data sets, each containing 10 haplotypes, were generated for each level of linkage. The mean values of the  $d_i$ 's, the number of sites where  $A_1$  is represented  $i$  times in the sample, are shown by the bars. Each thin line indicates the interval between the 2.5 and 97.5 percentiles for the distribution of a given  $d_i$ .

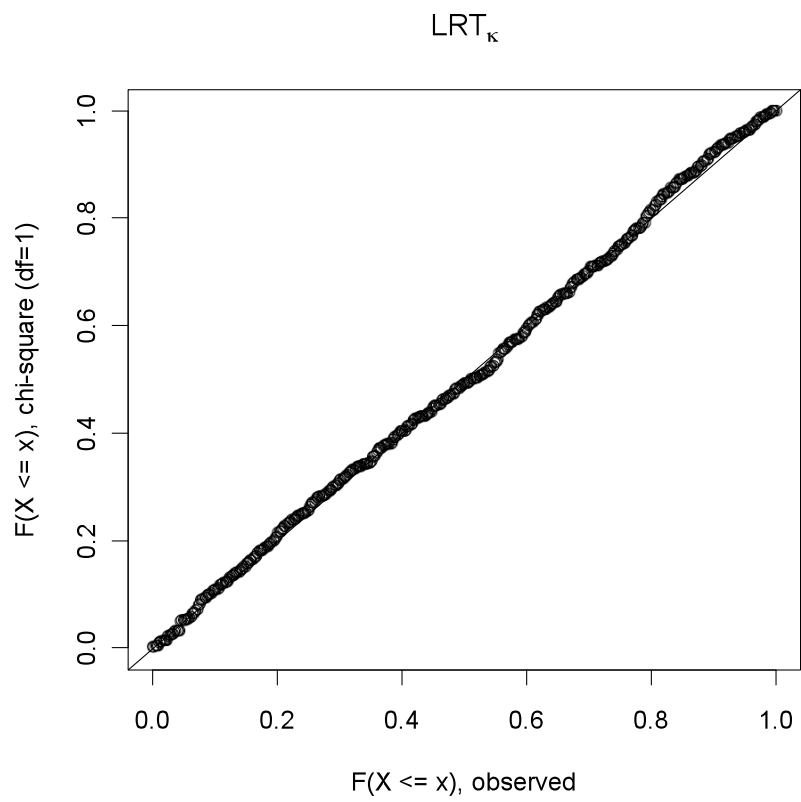
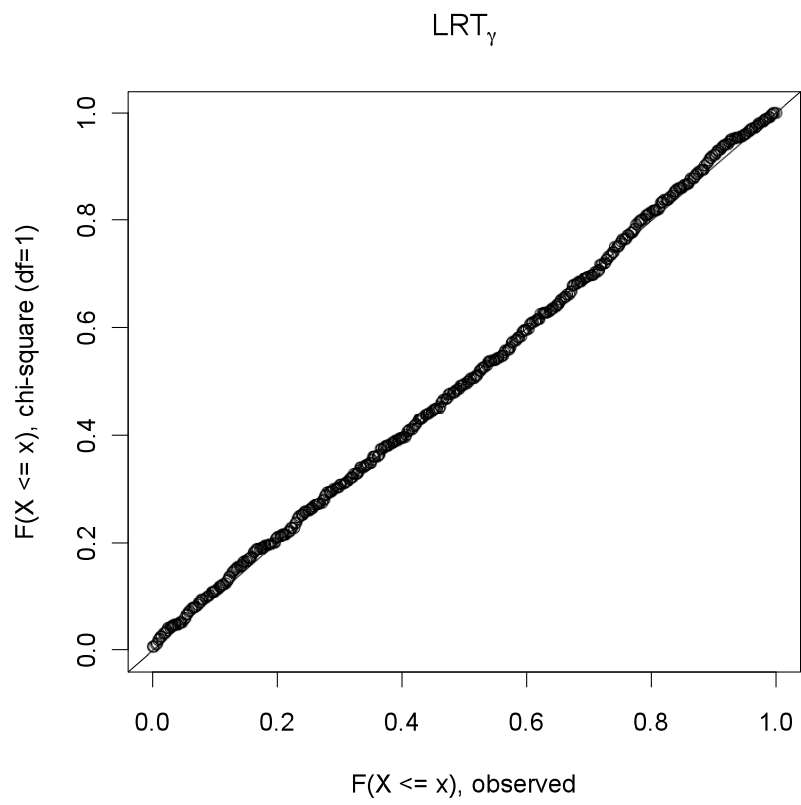
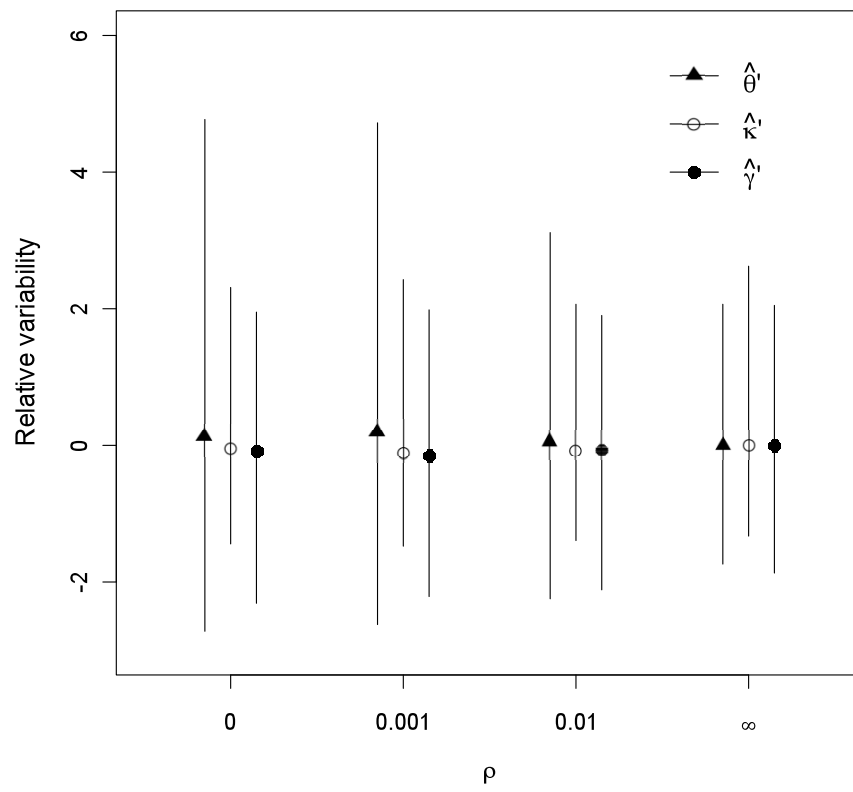
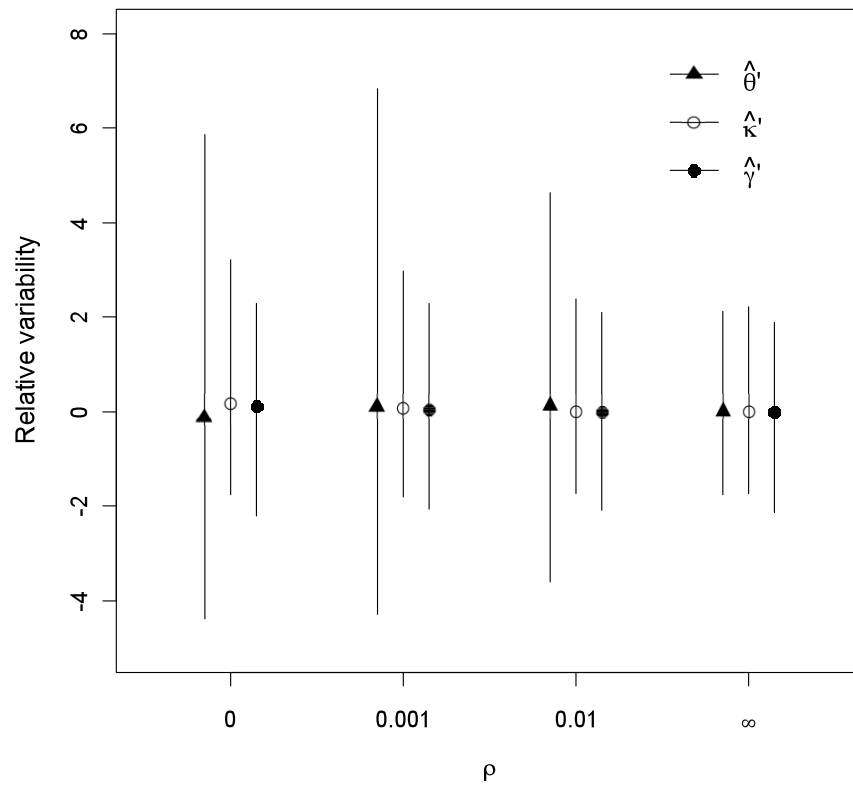


FIGURE S2.—Quantile-quantile plots of the distributions of  $LRT_\gamma$  and  $LRT_\kappa$  under complete linkage. Neutral evolution and complete linkage between sites were assumed. The size of the simulated region was set to 10Kb, with  $\theta = 0.01$  and  $\kappa = 1$ . Five hundred random samples were generated. We used the method of ZENG and CHARLESWORTH (2009) to analyze each of these samples. Likelihood ratio scores for the two tests,  $LRT_\gamma$  and  $LRT_\kappa$ , were obtained and compared to a chi squared distribution with one degree of freedom. The thin line is the 45-degree diagonal line.  $F(X \leq x)$  is the probability that the random variable  $X$  takes a value smaller than  $x$ . For the simulated data, this quantity was estimated by counting the number of likelihood ratio scores smaller than a given value.

**A****B**

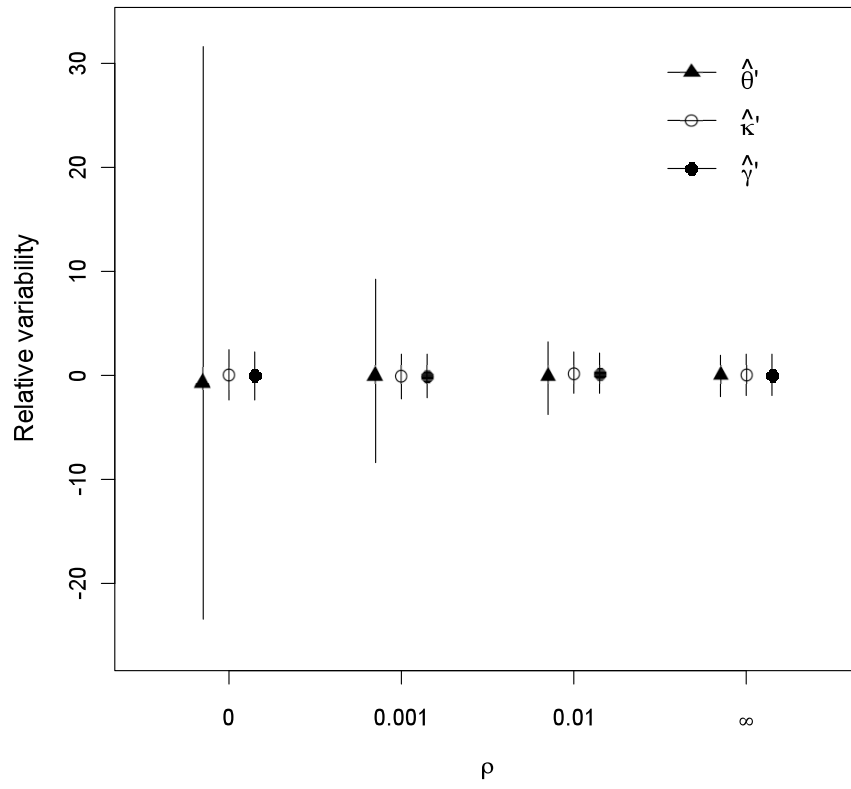
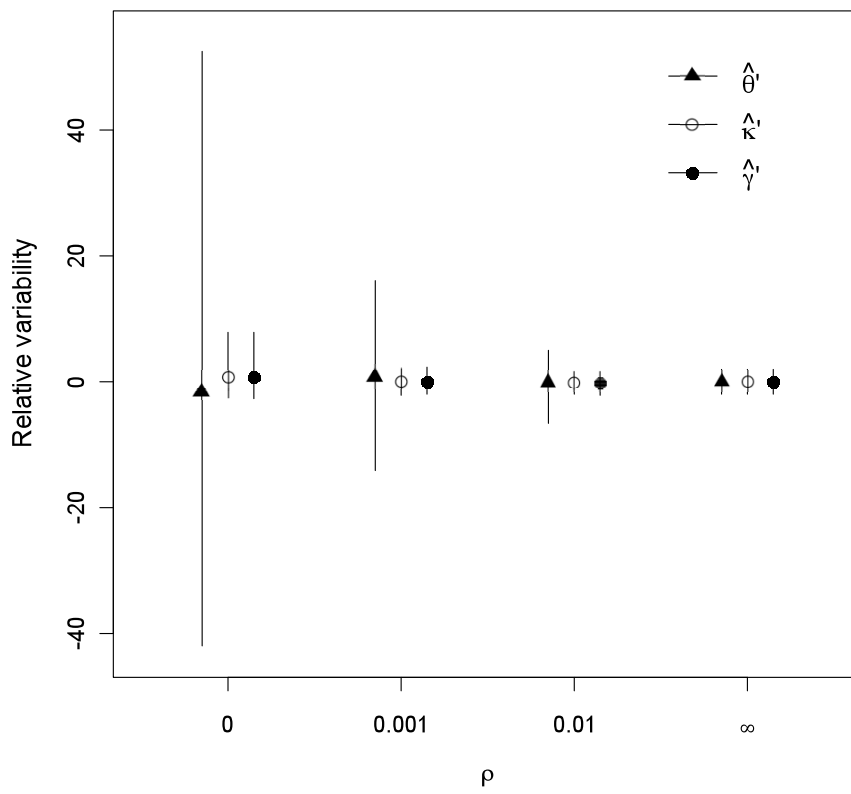
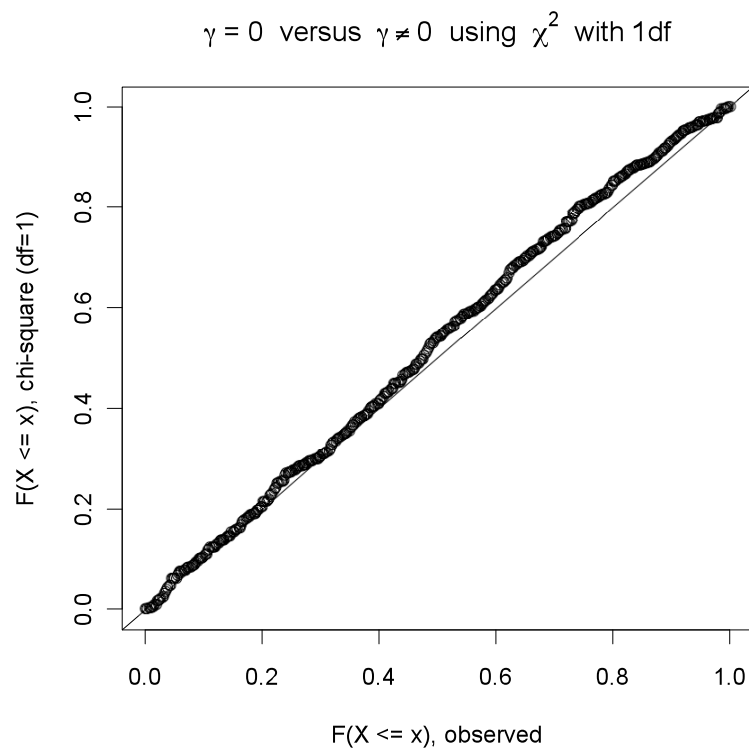
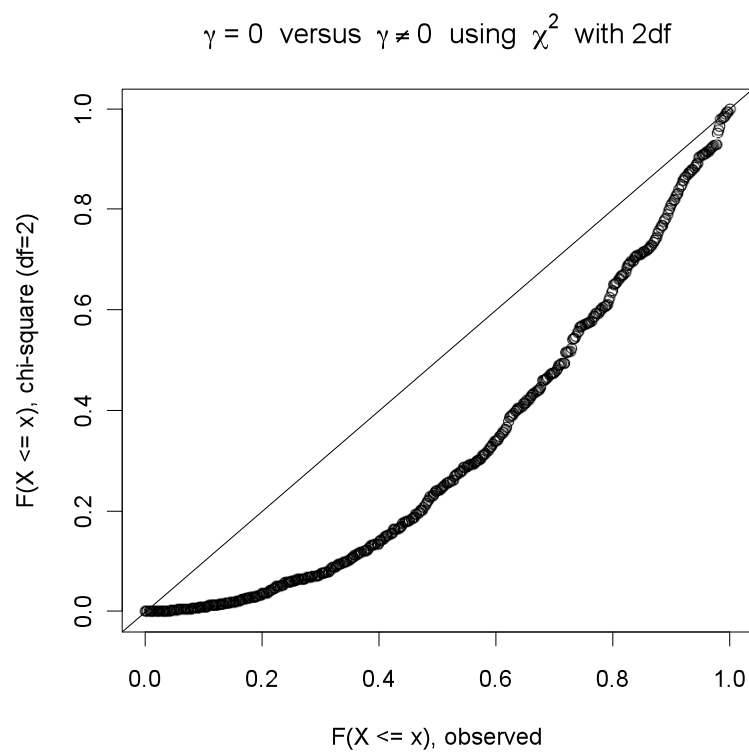
**C****D**



FIGURE S3.—The effects of linkage on the distributions of  $\hat{\theta}$ ,  $\hat{\kappa}$ , and  $\hat{\gamma}$ . The simulation model described in Figure 2 was used. The sample size was 15. The input parameters were: in (A),  $\theta = 0.02$ ,  $\kappa = 1$ , and  $L = 1\text{Kb}$ ; in (B),  $\theta = 0.05$ ,  $\kappa = 2$ , and  $L = 1\text{Kb}$ ; in (C),  $\theta = 0.02$ ,  $\kappa = 1$ , and  $L = 100\text{Kb}$ ; in (D),  $\theta = 0.05$ ,  $\kappa = 2$ , and  $L = 100\text{Kb}$ . The case of  $\rho = \infty$  corresponds to the free recombination assumption. For the ease of presentation, the  $\hat{\theta}$  values, for instance, were transformed using the formula  $\hat{\theta}' = [\hat{\theta} - E_{\rho}(\hat{\theta})] / SD_{\rho}(\hat{\theta})$ , where  $E_{\rho}(\hat{\theta})$  and  $SD_{\rho}(\hat{\theta})$  are the mean and standard deviation of  $\hat{\theta}$  obtained under free recombination, respectively; in the figure, the mean of  $\hat{\theta}'$  is indicated by  $\blacktriangle$ , and the intervals between the 2.5 and 97.5 percentiles are indicated by the solid lines. Similar procedures were applied to the values of  $\hat{\kappa}$  and  $\hat{\gamma}$ .

**A****B**

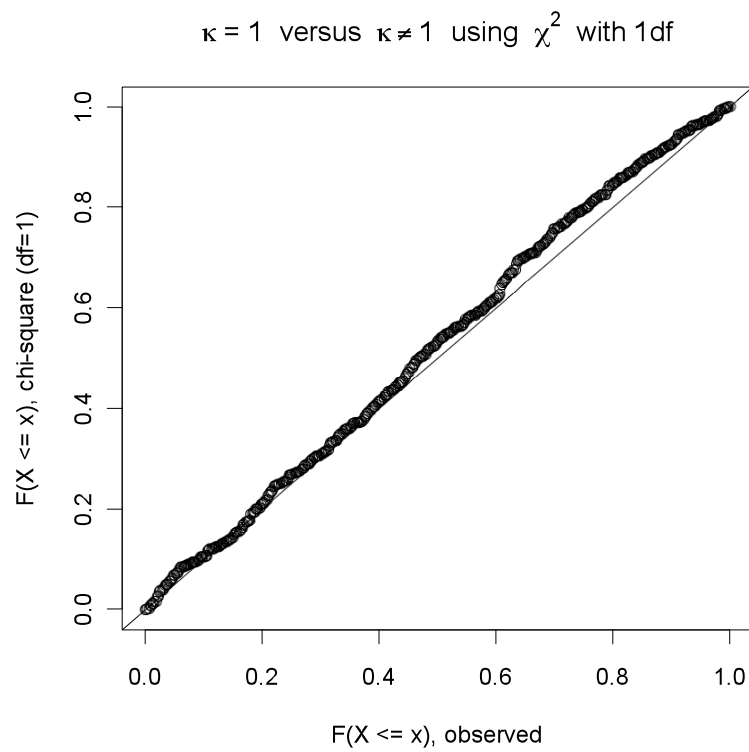
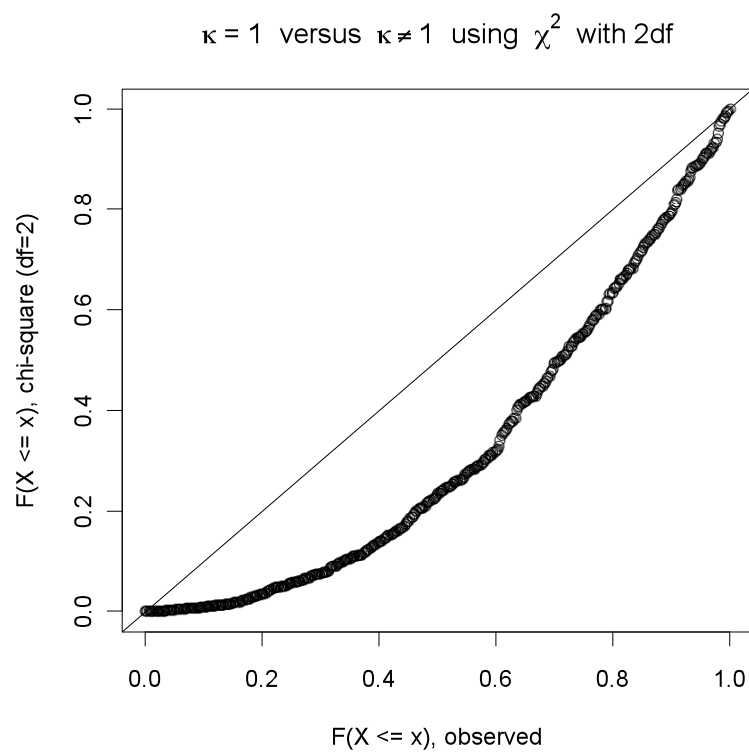
**C****D**

FIGURE S4.—Performing  $LRT_\gamma$  and  $LRT_\kappa$  using chi-squared distributions with one or two degrees of freedom in the presence of recent population expansion. The data were generated as in Table 3. Time was set to  $\tau = 0.04$ . The observed distributions of  $LRT_\gamma$  and  $LRT_\kappa$  were compared to chi-squared distributions with one or two degrees of freedom, respectively. We visualize the comparisons by using quantile-quantile plots as in Figure S2. If an observed distribution is conformed to an assumed distribution, the thick curve should fall on the 45-degree diagonal line (thin straight line). If the observed curve falls consistently above the diagonal line, it means performing the test using the assumed distribution is counter-conservative. On the other than, if the observed curve falls consistently below the diagonal line, it means that using the assumed distribution tends to produce conservative results.

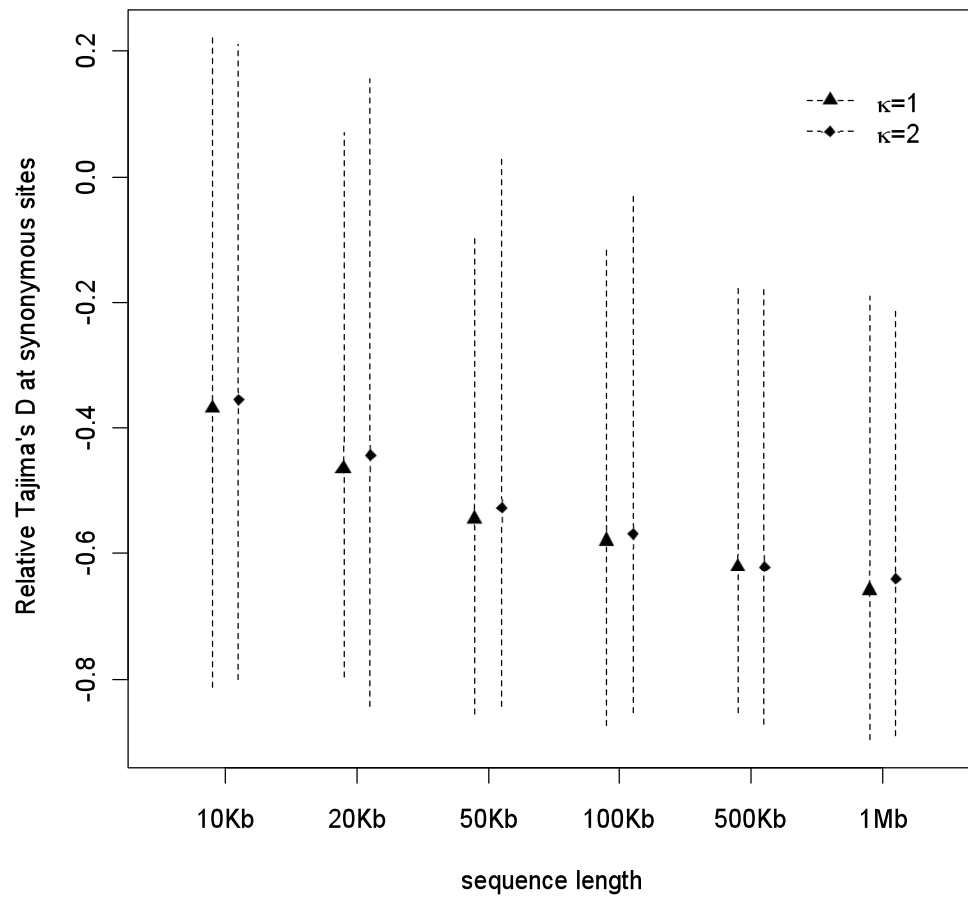


FIGURE S5.—The effects of mutational bias on the frequency spectrum at synonymous sites. The simulation procedures were the same as those used to obtain Figure 2. Complete linkage between sites was assumed.

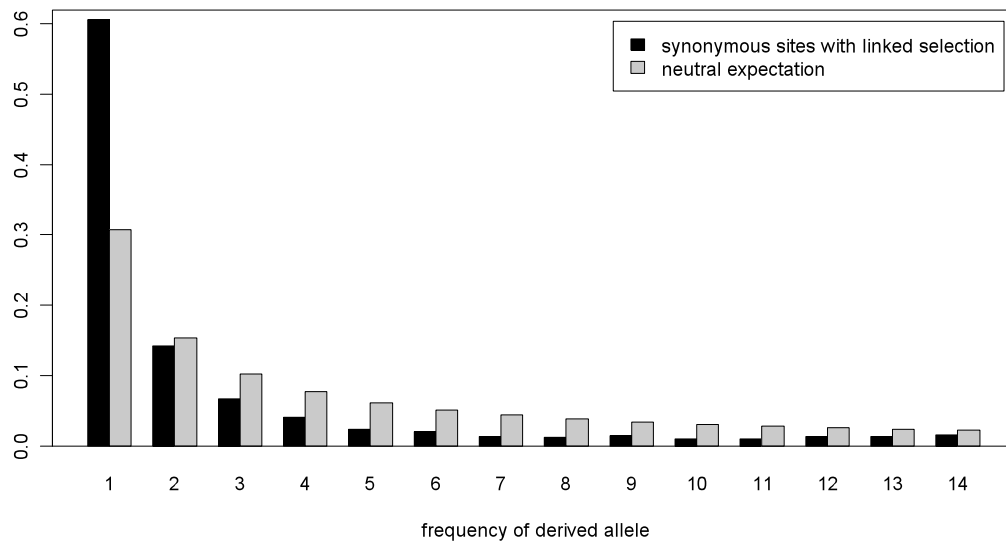
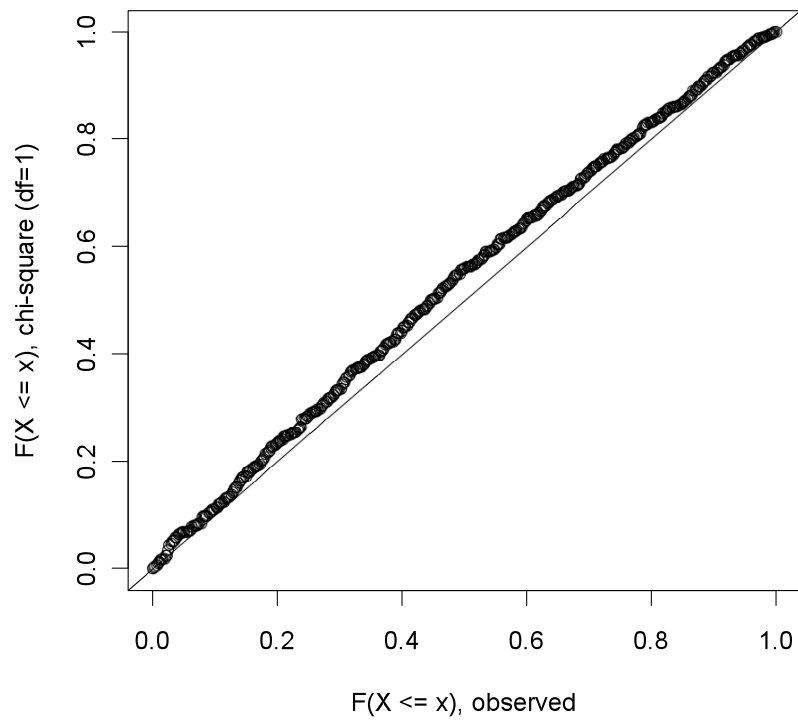
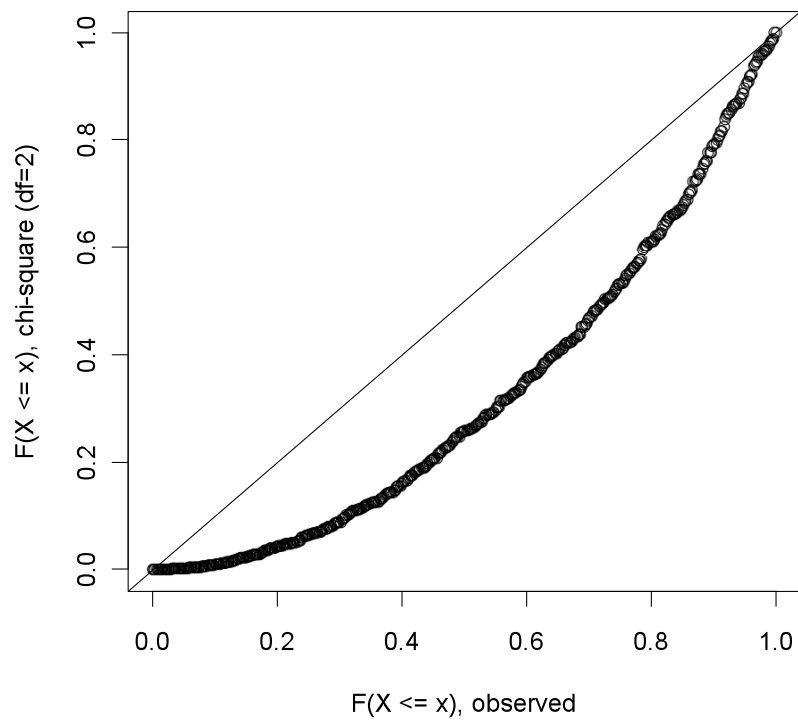
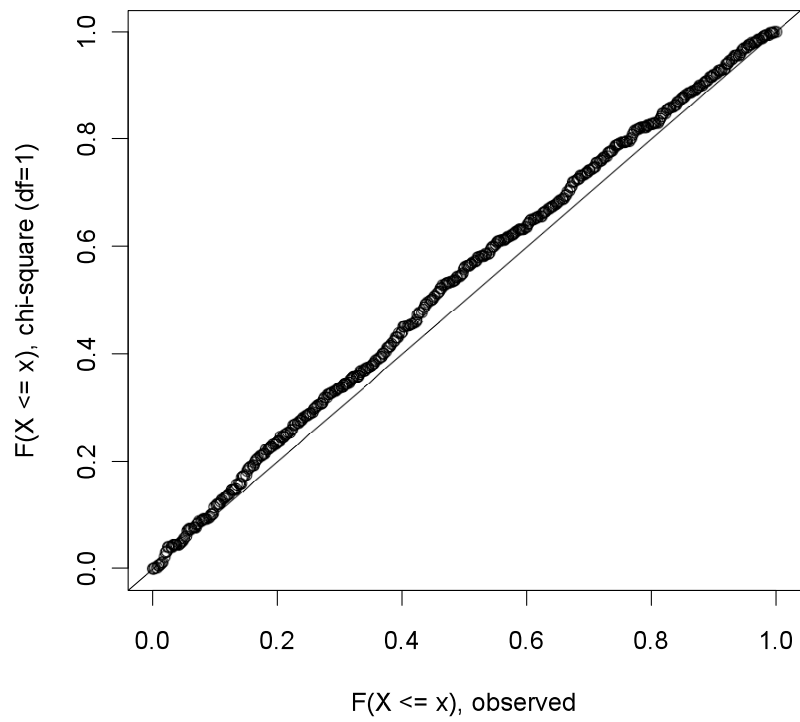


FIGURE S6.—The derived allele frequency spectrum at neutrally evolving synonymous sites in the presence of linked nonsynonymous sites. The simulation procedures were the same as those described in Figure 2. Parameter values were:  $\theta = 0.02$ ,  $\kappa = 1$ , and  $\rho = 0$ . The genealogy was tracked, and we used the ancestral sequence in the most recent common ancestor of the chromosomes in the sample to determine ancestral and derived states for each segregating site. The data presented here is based on 250 random samples.

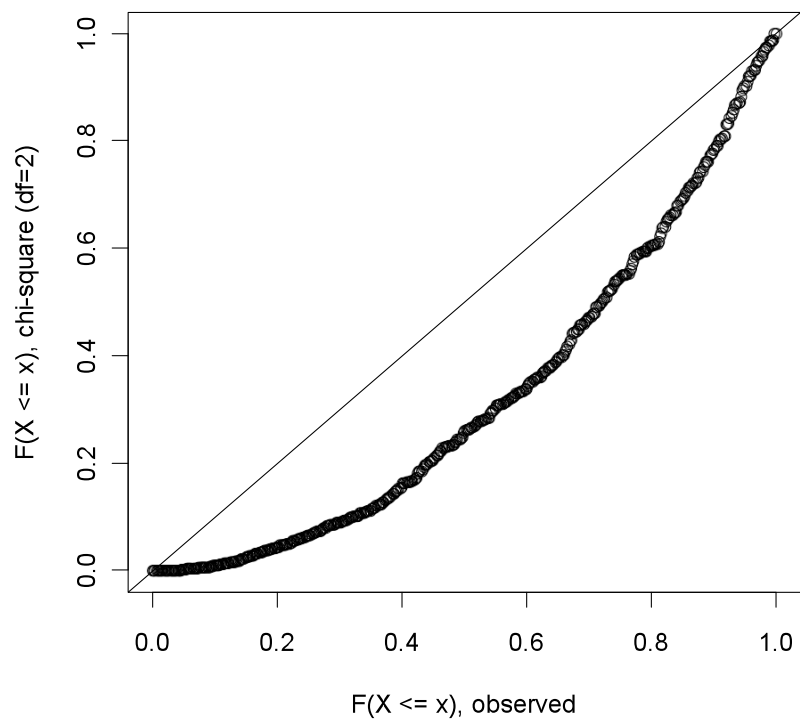


**A** $\gamma = 0$  versus  $\gamma \neq 0$  using  $\chi^2$  with 1df $\gamma = 0$  versus  $\gamma \neq 0$  using  $\chi^2$  with 2df

$\kappa = 1$  versus  $\kappa \neq 1$  using  $\chi^2$  with 1df

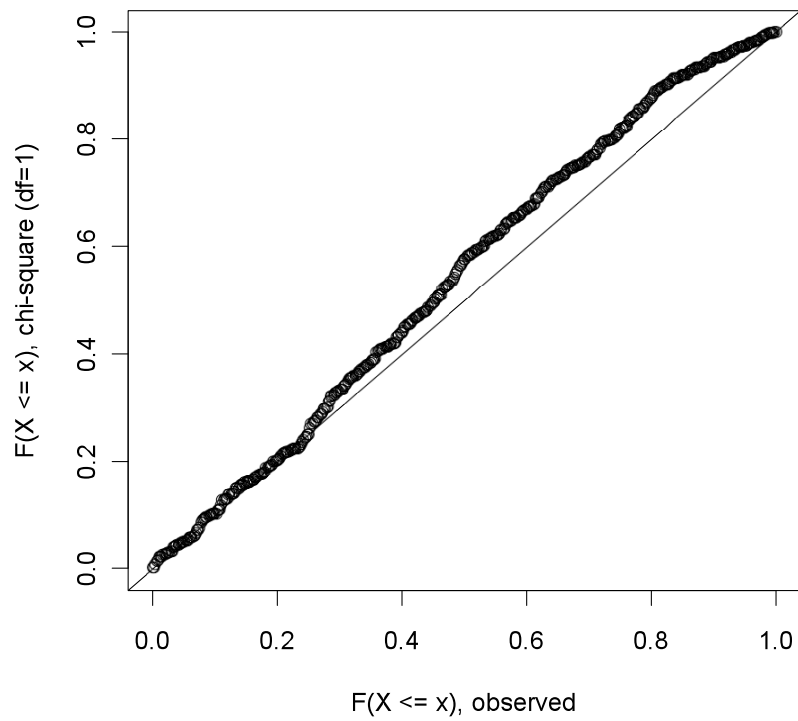


$\kappa = 1$  versus  $\kappa \neq 1$  using  $\chi^2$  with 2df

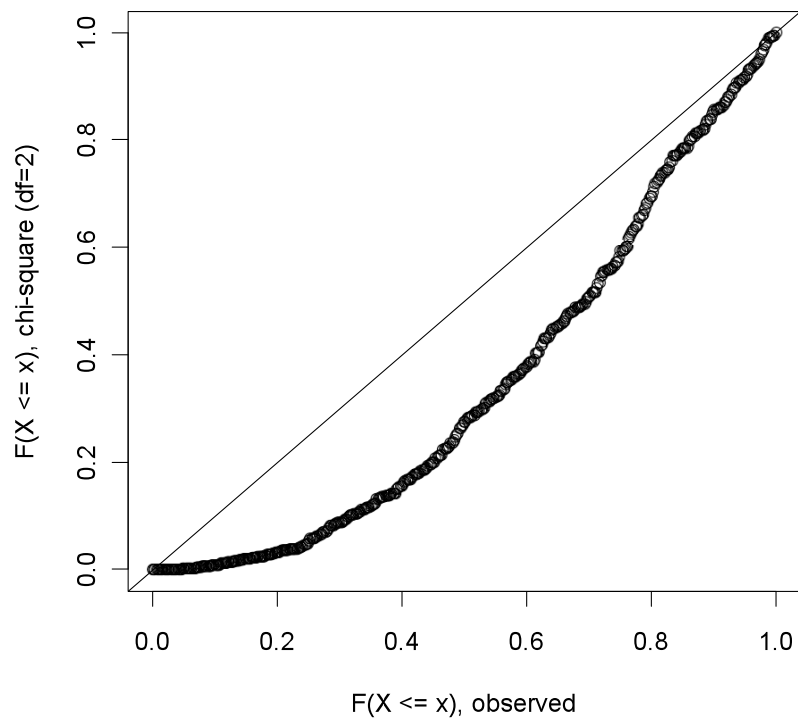


**B**

$\gamma = 0$  versus  $\gamma \neq 0$  using  $\chi^2$  with 1df



$\gamma = 0$  versus  $\gamma \neq 0$  using  $\chi^2$  with 2df



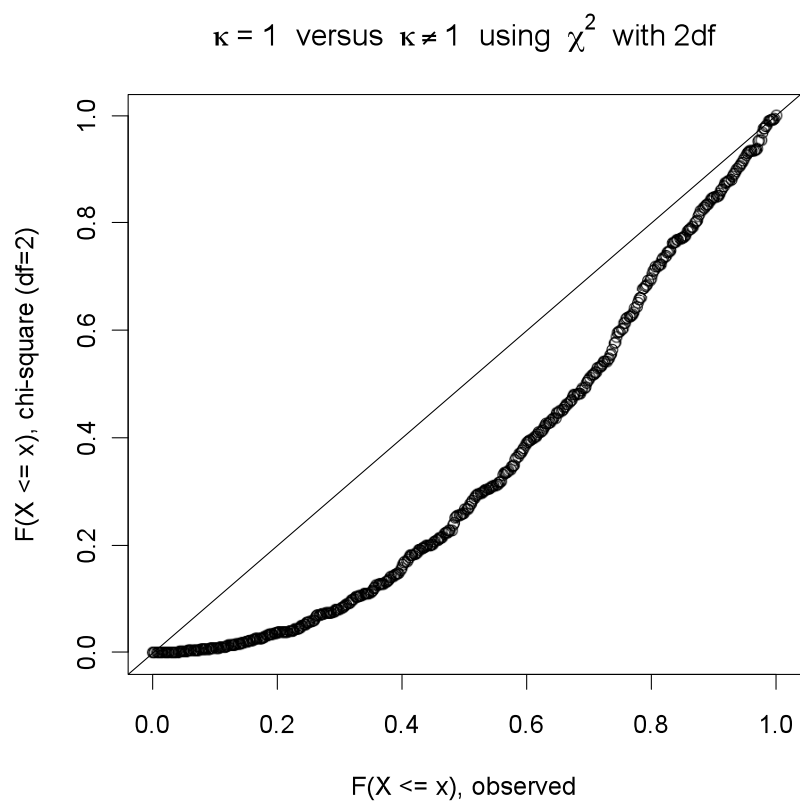
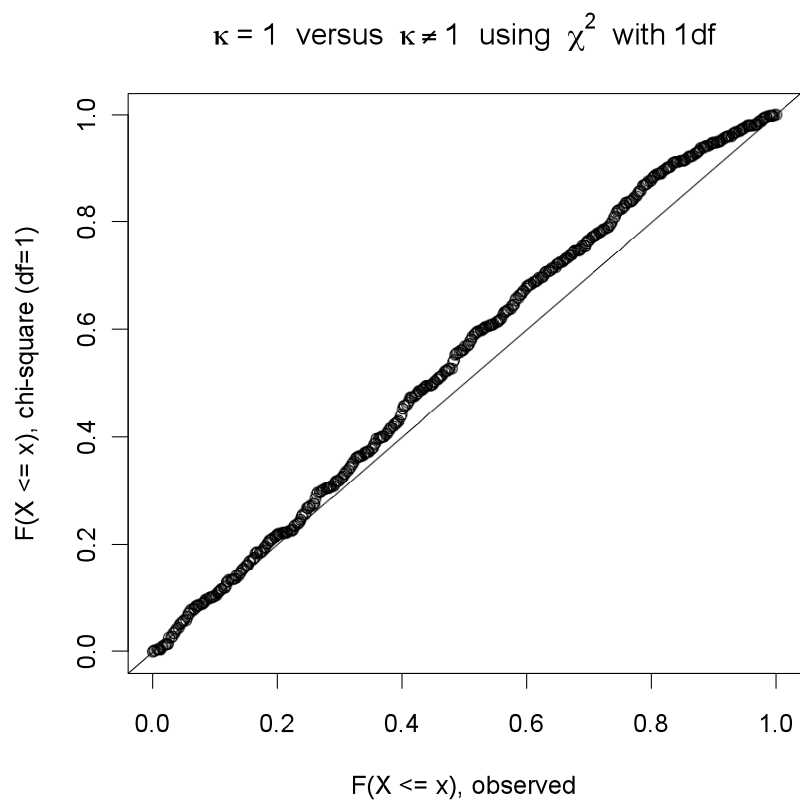


FIGURE S7.—Performing  $LRT_{\gamma}$  and  $LRT_{\kappa}$  using chi-squared distributions with one or two degrees of freedom in the presence of linked nonsynonymous sites. The data were generated as in Table 6. The parameter values used in all plots were  $\theta = 0.02$ ,  $\kappa = 1$ , and  $\rho = 0$ . In (A),  $L = 10\text{Kb}$ ; in (B),  $L = 50\text{Kb}$ . The observed distributions of  $LRT_{\gamma}$  and  $LRT_{\kappa}$  were compared to chi-squared distributions with one or two degrees of freedom, respectively. We visualize the comparisons by using quantile-quantile plots as in Figure S4.

### The effects of linkage on the method of ZENG and CHARLESWORTH (2009)

[illegible]

0.01	1	0.907	0.757	7.29	2.02	6.68	2.63	0.010	1.02	0.005
0.05	1	0.609	0.570	5.86	2.42	4.44	1.82	0.050	1.00	-0.003
0.04	2	0.191	0.000	6.81	1.40	99.60	99.40	0.020	2.03	0.007
0.04	4	0.089	0.000	6.41	2.20	100.00	100.00	0.010	4.11	0.014
0.05	2	0.117	0.000	6.60	2.40	99.80	99.60	0.025	2.03	0.010
0.05	4	0.122	0.000	5.20	1.80	100.00	100.00	0.012	4.08	0.011
0.06	2	0.000	0.000	9.98	4.07	100.00	100.00	0.032	2.03	0.011
0.06	4	0.054	0.000	7.83	2.21	100.00	100.00	0.015	4.13	0.024
0.07	2	0.001	0.000	9.24	4.42	100.00	100.00	0.036	2.04	0.017
0.07	4	0.008	0.000	8.60	3.00	100.00	100.00	0.018	4.06	0.008
0.08	2	0.021	0.000	9.02	3.41	100.00	100.00	0.041	2.03	0.012
0.08	4	0.000	0.000	11.62	5.81	100.00	100.00	0.021	4.14	0.030
$n = 15$ $L = 100\text{Kb}$										
0.02	1 ( $\rho = 0$ )	0.144	0.219	7.20	1.20	6.40	2.00	0.020	1.00	0.001
0.1	2 ( $\rho = 0$ )	0.000	0.000	28.00	16.80	100.00	100.00	0.050	2.04	0.021
0.1	2 ( $\rho = 0.001$ )	0.419	0.000	6.00	2.00	100.00	100.00	0.052	2.00	-0.000
0.1	2 ( $\rho = 0.01$ )	0.617	0.000	3.60	2.40	100.00	100.00	0.051	1.99	-0.008

---

The data generating procedures were as described in Table 2. In all cases,  $\rho = \phi = 0$ , except when  $n = 15$  and  $L = 100\text{Kb}$ .





0.42	7.9E-08	5.0E-07	3.0	2.6	33.4	0.2	0.2	1.02	0.004
0.45	1.0E-08	5.3E-07	4.0	3.6	36.6	1.6	1.6	1.03	0.000
0.48	2.3E-06	1.1E-06	1.8	1.8	34.2	0.2	0.2	1.02	-0.003
0.6	1.2E-06	9.2E-07	3.6	2.6	36.4	0.6	0.8	1.05	0.016
0.63	4.7E-08	1.7E-07	3.6	3.6	35.2	1.8	1.6	1.04	0.002
0.66	2.1E-03	3.4E-04	3.0	3.2	37.0	1.0	1.0	1.04	0.004
$\mathcal{N}_2/\mathcal{N}_1 = 0.1, \kappa = 2$									
0.42	1.2E-02	0	5.0	84.4	32.8	2.4	74.1	2.10	0.016
0.45	4.1E-03	0	3.6	81.0	31.8	1.2	71.0	2.09	0.009
0.48	9.4E-02	0	3.4	84.0	36.0	1.0	72.6	2.12	0.025
0.6	8.8E-06	0	4.4	77.0	36.8	2.4	65.4	2.14	0.022
0.66	1.5E-05	0	3.8	77.0	39.0	1.4	63.6	2.12	0.014

---

**TABLE S3****The effects of sampling on  $LRT_\gamma$  and  $LRT_\kappa$  in a subdivided population**

Sampling	$p$ value of KS test		Rejection rate (in %)		Mean MLE	
	$LRT_\gamma$	$LRT_\kappa$	$LRT_\gamma$	$LRT_\kappa$	$\hat{\kappa}$	$\hat{\gamma}$
Two-deme model, $4Nm = 0.5$ , $F_{ST} = 0.33$ , $\kappa = 1$						
[1, 14]	0.654	0.633	6.2	5.2	1.06	0.013
[2, 13]	0.622	0.706	5.0	4.6	1.07	0.026
[3, 12]	0.240	0.500	5.2	4.8	1.07	0.031
[4, 11]	0.391	0.290	5.2	4.8	1.06	0.027
Two-deme model, $4Nm = 0.5$ , $F_{ST} = 0.33$ , $\kappa = 2$						
[1, 14]	0.153	0	4.2	55.9	2.15	-0.003
[2, 13]	0.989	0	4.8	58.6	2.11	-0.007
[3, 12]	0.958	0	3.2	59.7	2.09	-0.014
[4, 11]	0.962	0	3.8	60.1	2.08	-0.013
[5, 10]	0.446	0	5.2	61.6	2.08	-0.015
[6, 9]	0.350	0	4.6	60.3	2.06	-0.020
[7, 8]	0.032	0	4.4	61.1	2.05	-0.022
Five-deme model, $4Nm = 0.5$ , $F_{ST} = 0.56$ , $\kappa = 1$						
[1, 14]	0.470	0.449	5.6	5.6	1.04	0.019
[2, 13]	0.201	0.323	4.6	5.4	1.04	0.018
[3, 12]	0.776	0.889	4.6	3.8	1.03	0.016
[4, 11]	0.137	0.114	2.4	3.0	1.03	0.016
Five-deme model, $4Nm = 0.5$ , $F_{ST} = 0.56$ , $\kappa = 2$						
[1, 14]	0.174	0	6.6	85.8	2.06	0.001
[2, 13]	0.043	0	5.6	89.6	2.06	0.008

[3, 12]	0.337	0	6.8	90.6	2.07	0.013
[4, 11]	0.115	0	4.8	93.0	2.08	0.021
[5, 10]	0.002	0	5.2	92.8	2.08	0.024
[6, 9]	0.001	0	4.4	93.0	2.08	0.025
[7, 8]	0.000	0	3.0	93.0	2.08	0.027

---

The simulation models are the same as those used in Table 4. The sampling scheme is described by  $[i, j]$ , where  $i$  and  $j$  give the number of chromosomes taken from the first and second demes, respectively ( $i + j = 15$ ).

**TABLE S4****The effects of selection at linked nonsynonymous sites on applying  $LRT_\gamma$  and  $LRT_\kappa$  to neutrally evolving synonymous sites.**

Input			$p$ value of KS test		Rejection rate (in %)					Mean MLE		
No. replicates	$\kappa$	$L$	$LRT_\gamma$	$LRT_\kappa$	$LRT_\gamma$	$LRT_\kappa$	$D_T$	$LRT_{\gamma^2}$	$LRT_{\kappa^2}$	$\hat{\theta}$	$\hat{\kappa}$	$\hat{\gamma}$
500	1	20Kb	0.490	0.282	7.0	7.2	26.4	2.8	3.0	0.0038	1.10	0.017
500	2	20Kb	$1.2 \times 10^{-5}$	0	8.4	40.6	22.6	2.8	25.8	0.0019	2.31	0.007
300	1	500Kb	0.158	0.162	14.0	13.0	62.1	1.0	1.0	0.0007	1.01	-0.008
200	2	500Kb	0.034	0	11.0	89.0	58.5	4.0	83.0	0.0004	2.07	-0.001
100	1	1Mb	0.301	0.291	10.0	10.0	74.0	4.0	4.0	0.0004	1.00	-0.016
100	2	1Mb	$9.3 \times 10^{-4}$	0	8.1	96.0	65.7	2.0	91.9	0.0003	2.05	0.006

The simulation procedures were identical to those used to obtain Figure 2. Complete linkage between sites was assumed.

Synonymous sites were assumed to be neutral. We set  $\kappa\theta = 0.02$ . Each random sample contained 15 chromosomes.